Technical Report 971

Use of Virtual Environment Training Technology for Individual Combat Simulation

William H. Levison and Richard W. Pew
BBN Systems and Technologies, Inc.

February 1993

United States Army Research Institute
for the Behavioral and Social Sciences

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Acting Director

Research accomplished under contract
for the Department of the Army

BBN Systems and Technologies, Inc.

Technical review by

Margaret S. Salter
Bob G. Witmer

Accesion For		
NTIS CRA&I	☒	
DTIC TAB	☐	
Unannounced	☐	
Justification		
By		
Distribution /		
Availability Codes		
Dist	Avail and / or Special	
A-1		

NOTICES

DISTRIBUTION: Primary distribution of this report has been made by ARI. Please address correspondence concerning distribution of reports to: U.S. Army Research Institute for the Behavioral and Social Sciences, ATTN: PERI-POX, 5001 Eisenhower Ave., Alexandria, Virginia 22333-5600.

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE	Form Approved OMB No 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	1993, February	Final Sep 91 - Apr 92

4. TITLE AND SUBTITLE
Use of Virtual Environment Training Technology for Individual Combat Simulation

5. FUNDING NUMBERS
N61339-91-C-0143
62785A
790
2111
C02

6. AUTHOR(S)
Levison, William H.; and Pew, Richard W.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
BBN Systems and Technologies, Inc.
10 Moulton Street
Cambridge, MA 02138

8. PERFORMING ORGANIZATION REPORT NUMBER
--

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Institute for the Behavioral and Social Sciences
ATTN: PERI-IF
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
ARI Technical Report 971

11. SUPPLEMENTARY NOTES
Performed under contract N61339-91-C-0143 with the Naval Training Systems Center. Contracting Officer's Representative, Bruce W. Knerr.

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE
--

13. ABSTRACT *(Maximum 200 words)*

For this research, 25 tasks and functions for dismounted soldiers were derived from Army training documents. Ten types of virtual environment technology were identified: visual display, visual sensing, auditory display, auditory sensing, haptic display, haptic sensing, whole-body movement, biomechanical articulation of Dismounted Infantry (DI) models, influence of physical condition on DI models, and physical condition of trainee. For each type of technology, up to three levels of capability were identified that represent anticipated availability in the near, intermediate (3-5 years), or far (more than 5 years) term. Subjective estimates were then made of the minimum level of technology required to support training of each of the tasks and functions. Although the technology that will be available in the near term does not appear to provide adequate training on all tasks and functions, it will nevertheless provide potential training and mission rehearsal benefits. Difficult problems yet to be resolved concern mission-specific training, urban and close-in operations, control and manipulation of weapons and equipment, and whole-body movement.

14. SUBJECT TERMS
Dismounted infantry
Virtual reality
Virtual environments
Simulation networking (SIMNET)
Distributed interactive simulation

15. NUMBER OF PAGES
92

16. PRICE CODE
--

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

NSN 7540-01-280-5500

Standard Form 298 (Rev 2-89)
Prescribed by ANSI Std Z39-18
298-102

Technical Report 971

Use of Virtual Environment Training Technology for Individual Combat Simulation

William H. Levison and Richard W. Pew
BBN Systems and Technologies, Inc.

STRICOM, Orlando Field Unit, Florida
Stephen L. Goldberg, Chief

Training Systems Research Division
Jack H. Hiller, Director

U.S. Army Research Institute for the Behavioral and Social Sciences
5001 Eisenhower Avenue, Alexandria, Virginia 22333-5600

Office, Deputy Chief of Staff for Personnel
Department of the Army

February 1993

Army Project Number
2Q162785A790

Human Performance Effectiveness
and Simulation

Approved for public release; distribution is unlimited.

FOREWORD

The Army has made a substantial commitment to Distributed Interactive Simulation (DIS) and the electronic battlefield for training, concept development, and test and evaluation. The current DIS training system, Simulation Networking (SIMNET), and the next generation system, the Close Combat Tactical Trainer (CCTT), provide effective training for soldiers fighting from vehicles, but are unable to do the same for individual dismounted soldiers. Virtual Environment Training Technology (VETT) has the potential to provide Individual Combat Simulations (ICS) for the electronic battlefield.

The potential of VETT for Individual Combat Simulations was explored as part of a cooperative effort with the Naval Training Systems Center (NTSC). The objective of the joint effort was to review the state of the art in VETT with respect to military training requirements and to identify research issues and needs.

The work described is part of a U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) research task entitled VIRTUE--Virtual Environments for Combat Training and Mission Rehearsal. The cooperative effort with NTSC is documented in a memorandum of agreement between NTSC and ARI entitled "Use of Virtual Environment Training Technology for Individual Combat Simulations" and dated September 1991.

EDGAR M. JOHNSON
Acting Director

ACKNOWLEDGMENTS

This report was prepared by BBN Systems and Technologies for the Naval Training Systems Center. The intent of the effort was to apply and extend previous work done for the Navy to the needs of the Army. David Fowlkes served as the contracting officer's representative for the Navy and Bruce Knerr served as representative of the U.S. Army Research Institute for the Behavioral and Social Sciences Field Unit at STRICOM, Orlando, Florida.

The author acknowledges the contribution of professor Nathaniel L. Durlach who, as consultant to this effort, provided highly constructive criticism of the work and helped flesh out some of the ideas presented in the report. The authors wish also to acknowledge their appreciation for the guidance and commentary provided by Dr. Knerr.

USE OF VIRTUAL ENVIRONMENT TRAINING TECHNOLOGY FOR INDIVIDUAL COMBAT SIMULATION

EXECUTIVE SUMMARY

Requirement:

This research reviews the state of the art in Virtual Environment Training Technology (VETT) from the perspective of requirements for individual combat simulation (ICS), i.e., the capability to insert the dismounted soldier into simulated combat for purposes of mission planning and rehearsal, mission-specific training, and combat proficiency training.

Procedure:

Research derived 25 tasks and functions for dismounted soldiers from Army Training and Evaluation Program documents. Ten types of virtual environment technology were identified: visual display, visual sensing, auditory display, auditory sensing, haptic display, haptic sensing, whole-body movement, biomechanical articulation of Dismounted Infantry (DI) models, influence of physical condition on DI models, and physical condition of trainee. For each type of technology, up to three levels of capability were identified that represent anticipated availability in the near, intermediate (3-5 years), or far (more than 5 years) term. Subjective estimates were then made of the minimum level of technology required to support training of each of the tasks and functions.

Findings:

Although the technology expected to be available in the near term does not appear to provide adequate training on all tasks and functions, there are potential training and mission rehearsal benefits to be obtained. Difficult problems to be resolved concern mission-specific training, urban and close-in operations, control and manipulation of weapons and equipment, and whole-body movement.

Utilization of Findings:

These findings, together with the results of a companion state-of-the-art assessment, provide the basis for a U.S. Army Research Institute for the Behavioral and Social Sciences research program to improve the Army's capability to provide effective, low-cost training for Special Operations Forces (SOF) and DI using ICS. Subsequent research reports will include a more detailed analysis of combat tasks, a review of the training and human performance research literature related to those task requirements to summarize requirements and identify research needs, and the establishment of a virtual environment research facility.

USE OF VIRTUAL ENVIRONMENT TRAINING TECHNOLOGY FOR INDIVIDUAL COMBAT SIMULATION

CONTENTS

CONTENTS (Continued)

LIST OF TABLES

USE OF VIRTUAL ENVIRONMENT TRAINING TECHNOLOGY FOR INDIVIDUAL COMBAT SIMULATION

1. INTRODUCTION

In June 1991 the Naval Training Systems Center (NTSC) entered into a contract with BBN Systems and Technologies, the Massachusetts Institute of Technology, Brandeis University, and some additional consultants to review the state of the art with respect to Virtual Environment Training Technology (VETT) and to develop a program plan and research agenda to realize the Navy's goals with respect to VETT.

The U.S. Army Research Institute Field Unit at STRICOM, Orlando, having a strong interest in VETT but different application priorities than the Navy, prepared an option to the NTSC contract that sought to take advantage of the work being accomplished for the Navy and to make it relevant to the Army's needs. This option was exercised in November 1992. This is the final report of that effort. The report is written to be a companion to the Navy volume (Durlach, Pew, Aviles, DiZio, and Zeltzer, 1992). Frequent reference is made to it and detail presented there is not repeated here. As indicated in the next section, the goals of this project have been to develop the potential application of VETT to the individual combat simulation.

1.1 Goals of the Project

This phase of the contract had four major goals:

Determine projected trends in capabilities and uses of individual combat simulations (ICS), with emphasis on training dismounted infantry.

Forecast the opportunities and problems associated with using Virtual Environment Training Technology (VETT) to integrate individual soldiers into these simulations for purposes of combat proficiency training, mission planning and rehearsal, and mission-specific training.

Identify major hardware and software requirements to allow effective utilization of VETT for ICS.

Specify the research tasks and facility requirements necessary to support ICS research.

Documents provided by the Army, particularly the description of the desired Close Combat Tactical Trainer (CCTT) and the Army Training and Evaluation Program for the Infantry Rifle Platoon

and Squad (Department of the Army, 1988), were used as the primary basis for assessing training requirements. A review of the state of the art of Virtual Environment (VE) Technology as it might be applied to Navy training needs (Durlach et al, 1992) was completed in the earlier phases of this contract; this review was used as the primary source of information concerning current and projected VE technology.

Because the emphasis of the study was on the application of virtual environment technology (VET) to combat training and because VET is largely concerned with the interface between the trainee and the simulated battlefield environment, our discussion of training needs and technology requirements focuses largely on the soldier/simulator interface. Some consideration is also given to other technology issues involved in providing an adequate simulation of the battlefield environment, such as iconic representation of the dismounted infantryman (DI).

We have concentrated on the training needs of dismounted infantry operating on the ground. Interface requirements for simulating activities relating to mounting and dismounting vehicles are considered, but we do not review interface issues related to soldiers who function primarily as mounted infantry (e.g., vehicle drivers and gunners).

Our baseline for specifying training and technology needs is derived largely from the requirements for initial implementation of the CCTT. The CCTT, which accommodates infantry, tank, and air support, is a component of the anticipated Combined Arms Tactical Trainer. Including the ICS as part of the battlefield simulation serves two goals: (1) providing realistic tactical training for the infantryman, and (2) providing a realistic battlefield environment for the training of other fighting elements (such as tank and close air support crews) whose actions are generally coordinated with those of infantry. Therefore, when designing a battlefield simulator to include DI, we must consider how the trainees are to be interfaced to the simulator, what modeling requirements are unique to the DI application, and how each dismounted infantryman - whether live or computer-based - is portrayed to the trainees and other observers of the simulated battle environment.

Initially, the DI trainees will consist of platoon leaders, forward observers, and squad leaders. The activities of other squad members will be modeled by the simulator in accordance with inputs provided by the squad leaders. Subsequent improvements and refinements of the simulated battle environment may accommodate additional trainees playing the roles of squad members.

This report deals mainly with the informational and technology requirements for providing adequate tactical training

to actual infantry personnel - initially, squad and platoon leaders. We also address issues related to the computer-generated portrayal of infantry personnel. Note that such portrayal is necessary even if all individual roles are played by live participants (a highly unlikely prospect), because each individual should be shown to all other individuals within line of sight as they would appear in a realistic battlefield environment.

1.2 Background

It is assumed that the personnel undergoing the types of training considered in this report have already been trained in the basics of soldiering and have had sufficient real or simulated exposure to the battle environment to understand at an intellectual level what they are supposed to do both in terms of goals and objectives and in terms of procedures and execution. The trainees are not learning to handle or dismantle weapons, dig trenches, and the like; rather, the soldiers are assumed to have acquired the necessary psychomotor skills, and the training is intended to integrate these individual skills into the tactical environment.

"Fidelity" in the context of the subsequent discussion refers to what we might term "training fidelity". In this sense, a high degree of simulator fidelity provides the trainees with informational quantities (and accepts command inputs) in such a manner as to cause the trainees to make assessments and perform activities in the same manner as they would in the operational tactical situation. A high degree of physical fidelity may or may not be required to provide a high degree of training fidelity, depending on the specifics of the situation.

This definition of simulator fidelity guides our thinking on simulator requirements as follows:

> Information provided need not be of a quality greater than that which can be discriminated by the trainee. Thus, human performance limitations may serve in some cases as a limit to the physical fidelity requirements of the simulator.
>
> Except for situations in which enhanced or artificial cues are introduced to speed up the learning process, training-relevant information provided to the trainee should be as good as, but no better than, the information that would be provided under the battlefield conditions represented in the simulation. For example, if distance, weather conditions, tactical smoke, time of day, or the soldier's degraded physical condition would realistically prevent the trainee from seeing an object clearly, and the course of the battle is influenced by how well the soldier would see that object, then the object should be presented with the same lack of

clarity as it would appear to have in the operational situation.

Emphasis should be on information transfer that maintains proper cause-and-effect relationships and not on physical mechanisms. For training procedural skills and for initial mission rehearsal, it may be sufficient for the squad leaders to command movement of their (computer-modeled) "troops" by means of a joystick, provided these troops move in a manner consistent with the simulated battlefield environment.

If an aspect of the battlefield environment is not expected to influence the soldier's behavior or state of knowledge in the operational situation, the degree of physical simulation fidelity is unimportant. For example, the time to load a weapon (and the possibility of running out of shells) may be important for tactical training and mission planning, but physical manipulation of the object (whether real or simulated) is unimportant for the types of training considered in this review.

Consistent with the above comments, all references to "informational requirements" imply information of the same quality that would be provided in the operational environment. For example, when we specify that the soldier/simulator interface allow the soldier to assess the suitability of the terrain for mechanized troop movement, we do not necessarily mean that the soldier be given information in a manner that allows an accurate assessment, but rather that the assessment (however poor) would be of the nature and accuracy achieved in the actual battlefield environment.

Three documents provided the bulk of the source material for this report: the above-mentioned review of VETT (Durlach et al. 1992), a document describing a mission training plan for infantry (Department of the Army, 1988), and the system specification for the Close Combat Tactical Trainer (Department of the Navy, 1991). Complete citations are given in Chapter 8.

1.3 Organization of the Report

The following chapter reviews the three purposes of training and provides a summary of the interface requirements for individual combat simulations. An overview of the VE technology relevant to combat training is provided in Chapter 3. Chapter 4 discusses the training requirements in more detail and suggests the type of VETT appropriate to meeting these requirements. Recommendations for technology development and behavioral research are presented in the following two chapters, and Chapter 7 summarizes the results of this report.

2. Interface Requirements for Individual Combat Simulations

This chapter reviews the soldier/simulator interface requirements for training the individual soldier. These requirements are discussed in terms of information-transfer needs rather than detailed technical specifications. We begin with a brief review of the three general types (or "purposes") of training contemplated: combat proficiency training, mission planning and rehearsal, and mission-specific training. Section 2.2 then presents a review of the top-level interface requirements for the various sensory modalities. Detailed discussion of these requirements is provided in Chapter 4.

2.1 Purposes of Training

Three generic training purposes are considered: combat proficiency training (CPT), mission planning and rehearsal (MPR), and mission-specific training (MST). Differences in terms of the nature and quality of the soldier/simulator interface imposed by these different training purposes are noted.

2.1.1 Combat Proficiency Training (CPT)

The primary purpose of combat proficiency training is to train combat units in the execution of various tactical missions. The emphasis is on operating as a unit, not on training or improving individual fighting skills. It is sufficient that the simulation not degrade individual fighting skills.

Tasks tend to be generic in nature. The simulated battle environment must be realistic, but not necessarily a replica of a specific battle situation. For example, to provide training in the techniques of capturing bridges, a series of bridges may be simulated in relatively close proximity to one another to allow the trainees to proceed rapidly from one bridge to the next. These bridges must be configured in a manner representative of real-world bridges, but they need not be replicas of existing bridges, and, even if the individual bridges are drawn from the database of actual bridges, their locations relative to one another may be substantially modified to speed up the training process.

In this report we consider CPT as representing a limited training regime in which the emphasis is on training the behaviors required of the realistic tactical environments, including the interactive coordinating behaviors associated with effective team performance. To meet this training goal the ICS must allow the trainee to perform a realistic assessment of the battle situation and to be shown a realistic outcome as a consequence of actions taken, and the simulation must show critical events unfolding in real time. Accurate representation

of the physical interaction of the soldier with the battlefield environment is less important.

Non-critical events may be omitted or performed in fast time. For example, in the hypothetical example offered above, the action of traveling from one captured bridge to the vicinity of the next bridge may be omitted or represented in fast time if the training is focussed strictly on the "end game" of actually securing the bridge.

To the extent that CPT goes beyond procedural training to include a more realistic exposure to the physical nature of the mission, the requirements for this type of training become identical to those for mission-specific training as discussed in Section 2.1.3.

2.1.2 Mission Planning and Rehearsal (MPR)

Mission planning, rehearsal, and training are performed with respect to achieving a specific real-world mission and often involve special operations forces. In this situation the simulation will, to the extent allowed by the available information, portray actual terrain, buildings, enemy force concentrations, etc.

One can envision the following stages proceeding from initial planning to final training in which increasing levels of simulator sophistication and complexity are required:

Using computerized models of all important elements of the battle environment (assuming such are available), perform the initial planning exercise using only a computer to explore the possible outcomes of various options and to begin to narrow the choices.

Introduce live players into the simulation to provide a more realistic simulation of mission progress and to allow individuals directly concerned with executing the mission to influence the mission plan.

Begin rehearsing the mission for the purposes of refining the plan and to provide some initial familiarization with the procedural and coordination skills required (e.g. when and where to perform what tasks).

Once the plan is complete, train the participants in all aspects of the mission with as much realism as is available and needed for training purposes. Ideally, rehearse the mission until the desired training goals have been met.

In keeping with current Army thinking, these four steps are combined into two training purposes: mission planning and

rehearsal (MPR), and mission-specific training (MST). We consider MPR to encompass the first three steps. That is, rehearsal will be considered to be an integral part of a planning process that also provides some procedural training for the mission.

As in the case of CPT, training at this stage need not replicate the detailed physical aspects of the activities involved in carrying out the mission. Furthermore, it may be possible to perform in fast time those portions of the mission that have been well practiced in the past and for which the timing and outcome can be predicted with a high degree of accuracy. To some extent, however, MPR will place greater demands on the simulator than does CPT because of the need to yield an accurate statistical prediction of mission outcome.

2.1.3 Mission-Specific Training (MST)

As in the case of MPR, mission-specific training (MST) pertains to specific operational missions and therefore requires relatively high validity with respect to simulating the details of the anticipated battle area. Unlike MPR, however, MST is intended to provide the trainees with a physical "feel" for the engagement as well as to reinforce the procedural training that may have been provided previously. In addition, it serves as a comprehensive check on the efficacy of the mission plan.

To meet the more demanding goals imposed at this stage, MST must replicate the entire mission in real time and, to the extent feasible and necessary for training purposes, replicate the physical aspects of the soldier's interaction with the environment.

2.2 Summary of Interface Requirements

In the remainder of this chapter we discuss the requirements of the soldier/simulator interface relevant to the goals of combat proficiency training, mission planning and rehearsal, and mission-specific training.

The discussion of interface requirements is organized largely in terms of sensory modality. Three primary classes of modality are considered: visual, auditory, and haptic. Haptic includes modalities related to physical contact and the sense of whole-body motion. Ways of accounting for the performance limitations that would be imposed by physical stress are discussed under the category "other requirements". A top-level overview is provided in this section; Chapter 4 discusses the interface requirements in more detail.

Table 1 provides a relatively high-level summary of the ICS interface requirements, organized in terms of sensory modality as

indicated above. The senses related most directly to detection and recognition of chemical substances - smell and taste - are omitted from discussion because, to the best of our knowledge, they have not been seriously addressed in the development of VE and simulator technologies.

2.2.1 Derivation of Tasks and Functions and Assignment to Modalities

The tasks and functions shown for each category were derived largely from an assessment of the informational requirements imposed by the military operational tasks described in the ARTEP for infantry rifle squad and platoon. Tasks and functions were assigned to modalities largely on the basis of physical requirements imposed by the operational world and, in some cases, by the expected limitations of near-term simulators. No significance is implied by the order in which tasks and functions appear within a given top-level category.

2.2.2 Applicability of Tasks and Functions for Particular Training Purposes within Modalities

The relative needs for various information and control capabilities are shown for CPT, MPR, and MST, where the code 0 indicates that the capability is not needed for the particular training purpose, 1 indicates that the capability is desired but not necessary for training, and 2 reflects the judgement that the capability is necessary for the particular training purpose. The requirements shown in Table 1 are stated from the point of view of the trainee in terms of what procedures they need to perform and what commands they need to execute, rather than what the simulator has to do (i.e., "perform navigation", rather than "provide navigational information").

Assignment of Code 2 ("capability necessary") was made on one of the following bases:

The task/function must be performed in order to accomplish one or more critical military operations, and performance requires the sensory modality indicated. All of the tasks/functions shown in Table 1a, Visual Interface Requirements, fall in this category.

The task/function must be performed, and the limitations of anticipated near-term simulation capabilities dictate the sensory modality shown. The task "command simulated troops" (Table 1c) is the clearest example of this consideration. The squad leader must be able to command troops, and this would be done in the operational world by voice and visual signalling. However, because the "troops" to be commanded in the CCTT will be computerized models of soldiers, and

Table 1

Summary of Informational Requirements for ICS Interface

	Training Purpose		
	CPT	**MPR**	**MST**
a) Visual Interface Requirements			
Perform navigation.	2	2	2
Assess physical environment with respect to cover and concealment.	2	2	2
Assess suitability of terrain and other aspects of physical environment for troop and vehicle movement.	2	2	2
Assess likelihood of enemy travel paths.	2	2	2
Conduct operations in an urban environment.	2*	2*	2*
Conduct non-urban close-in operations.	2*	2*	2*
Determine status (location, movement, activity, condition) of friendly troops.	2	2	2
Determine status of enemy forces.	2	2	2
Determine status of friendly and enemy vehicles.	2	2	2
Distinguish/Identify friendly and enemy forces.	2	2	2
Assess status and effects of weapons.	2	2	2
Recognize and assess status of various battle-field details (barbed wire, obstacles, etc.)	2	2	2

* Required for a limited set of situations

Training Purposes:
CPT = Combat Proficiency Training
MPR = Mission Planning and Rehearsal
MST = Mission Specific Training

Need for Capability:
0 = capability not needed
1 = capability desirable but not necessary
2 = capability necessary

Table 1 (Continued)

Summary of Informational Requirements for ICS Interface

	Training Purpose		
	CPT	MPR	MST
Read navigation aids.	2	2	2
Command troops via hand and arm signals.	0	0	1
b) Auditory Interface Requirements			
Communication among trainees.	2	2	2
Command simulated troops.	0	0	1
Conduct operations in an urban environment.	2*	2*	2*
Conduct non-urban close-in operations.	2*	2*	2*
Determine status of friendly and enemy vehicles.	1	1	2
Assess status and effects of weapons.	1	1	1
Be subjected to disruptive effects of battlefield noise.	1	1	2
c) Haptic and Other Sensory Interface Requirements			
Command simulated troops.	2	2	2
Receive feedback concerning physical condition of troops.	0	0	1
Perform control and manipulation of weapons, other equipment, objects.	2	2	2

* Required for a limited set of situations

Training Purposes:
CPT = Combat Proficiency Training
MPR = Mission Planning and Rehearsal
MST = Mission Specific Training

Need for Capability:
0 = capability not needed
1 = capability desirable but not necessary
2 = capability necessary

Table 1 (Concluded)

Summary of Informational Requirements for ICS Interface

	Training Purpose		
	CPT	**MPR**	**MST**
Receive realistic feedback from control and manipulation of weapons, equipment, objects.	1	0	1
Assess battlefield conditions.	1	0	1
Be subjected to physical and cognitive limitations of protective and other gear.	1	2	2
Perceive a sense of body movement.	0	0	1
d) Other Requirements			
Perceive effects of degraded physical condition of simulated DI.	1	2	2
Suffer physical and cognitive limitations on performance due to degraded physical condition.	1	2	2

---

* Required for a limited set of situations

Training Purposes:
- CPT = Combat Proficiency Training
- MPR = Mission Planning and Rehearsal
- MST = Mission Specific Training

Need for Capability:
- 0 = capability not needed
- 1 = capability desirable but not necessary
- 2 = capability necessary

because voice and visual recognition by the simulator is not anticipated in the near term, the (live) soldier must use a joystick or some other physical control device to indicate intended actions.

The task/function must be performed, it can be performed using another modality, but it is highly desirable for training purposes that the capability also be provided in the modality indicated. The one example of this case is the use of auditory information in determining the status of friendly and enemy vehicles for MST.

Carrying out the (simulated) military operation can be physically accomplished without the capability indicated, but it is our subjective judgement that realistic training requires that capability be provided. The items in Table 1d for MPR and MST, for example, fall in this category.

Code 1 (capability desirable but not necessary) indicates that performance of critical (simulated) military tasks does not require the capability in the modality shown, but that, in our judgement, training will be materially enhanced if the capability is provided. Similarly, Code 0 reflects our judgement that the anticipated training benefit of providing the capability in the indicated modality will be insufficient for the training purpose to warrant provision of this capability.

2.2.3 Special Considerations

Table 1 is intended to address the question of whether or not a given informational or control capability is needed. For the most part we have attempted to consider the degree of sophistication and complexity required of the simulator to meet training needs as an issue separate from what the training needs are. Technology requirements are discussed in Chapter 4.

One critical function of the soldier/simulator interface is to provide the soldier with a realistic assessment of the battle situation to allow planning, execution, and post-hoc evaluation of maneuvers, consistent with the degree to which such assessment could accurately be performed in the operational situation. Information of this sort is gained largely through the visual sense plus voice communication. Consequently, much of the discussion in this report concerns the presentation of visual information. To accommodate an adequate range of combat activities and environments, the ICS should be capable of presenting views as they would be seen from the standing, kneeling, or prone positions using unaided viewing, binocular viewing, or night vision equipment as appropriate.

To some extent, our assessment of simulator interface requirements is driven by the current and projected state of the

technology according to reasonable growth paths for available technology. For example, there are in principle at least three ways to allow the squad leader to command simulated "troops": (1) provide commands by voice and by arm and hand signals, (2) enter commands to the computer through a keyboard and/or switch panel, and (3) manipulate a joystick. If computer recognition of speech and visual signals were readily available, reliable, and cost effective, we would consider this mode of information transfer to be necessary and most likely not even consider the other two modes. Given that such technology requires further development, we consider a direct mechanical command interface to be necessary and speech recognition to be a desirable option, and even then only for the more intensive training modes.

2.2.4 Visual Interface Requirements

As expected, the largest set of requirements is associated with the visual interfaces. Furthermore, all but one of the visual capabilities relating to perceptual requirements at the level shown in Table 1a are considered necessary for all three training purposes. This conclusion is based on the judgement that proper training at any level requires that the trainee receive whatever information is needed and available to allow realistic assessment of the battle situation. Without this capability, training of unit tactical skills, refinement of mission plans, and mission training would be severely compromised. The level of technology needed to satisfy some of the visual information requirements differs according to training purpose as discussed in this report.

The two tasks in Table 1a that relate to rapidly-unfolding activities in close proximity to the soldier (operations in urban and other close-in environments) are broader in scope than the other entities, which focus more on specific informational requirements. These entries are included because, as we discuss later in the sections on VETT requirements, operations of this sort that require the soldier to make relatively large and rapid head movements place a higher demand on the VE technology needed to supply the visual information adequate for training.

2.2.5 Auditory Interface Requirements

To the extent that trainees interact with each other, voice communication (largely radio/telephone) will be the primary means of communication. For the near future, in which squad members are largely represented by computer models in the ICS, the (live) squad leaders will need to communicate directly with the simulator to command their "troops". Speech recognition capabilities are considered unnecessary for the more procedurally-oriented CPT and MPR kinds of training, provided that alternative interface mechanisms exist (e.g., joystick). Speech recognition is considered desirable for MST where it is

more important for the trainees to behave in an operationally realistic manner.

Auditory capabilities required for allowing the soldier to properly assess the battle environment from other than the spoken word are necessary for all three training purposes. This includes the capabilities to identify the event causing a sound and to assess its direction and distance. Auditory events associated more with "realism" than with information transfer are not considered necessary for CPT.

In addition to providing useful information, battlefield sounds will often have a deleterious influence on execution of the mission because of general stress effects or through interference with voice communications. Because such effects may influence the probability of mission success, it is important that such effects be accounted for in MST to allow assessment of the efficacy of the mission plan and to train the soldiers to operate effectively in such an environment.

The trainee must, of course, have some means of commanding troop movement. In the absence of speech recognition, a control interface relying on physical manipulation of some device (e.g., joystick) is necessary for all three training purposes. Similarly, the trainee must have the capability to operate weapons and other equipment, and to perhaps manipulate objects. (Again, the degree of sophistication of the soldier/simulator interface is a separate issue discussed later in this report.)

2.2.6 Haptic and Other Sensory Interface Requirements

MPR and MST may occasionally involve long-term operations where the progress (and possibly outcome) of the battle will be influenced by the physical condition of the troops. While information relating to the DI's fighting capability will most likely be provided largely through the observed appearance and behavior of the DI icon, it is considered desirable for MPR and MST to provide an additional "feel" for the DI's physical state through the control interface.

Physical feedback from weapons and other equipment that would be manipulated in actual battle is deemed less important for the largely tactical training considered in this report. (Recall that the soldier is assumed to have been previously trained in the operation of a particular piece of equipment and therefore has an idea of the "feel" of its operation.) Accordingly, feedback through the haptic sense is considered unneeded for planning and, at most, desirable but not necessary for CPT and MST. Instead, feedback needed to assess the status and orientation of the weapon will be provided through the visual and auditory senses.

There are circumstances when the sense of touch is useful for assessing battlefield conditions (e.g., test the surface conditions for sufficient weight-bearing capacity to serve as a landing pad). This capability would be useful for CPT and MST but is not necessary for mission planning purposes.

One aspect of the soldier/simulator interface that has been included in this category is the realistic simulation of the deleterious effects of the protective clothing worn in a suspected NBC environment, plus encumbrances imposed by other pieces of equipment or objects that the soldier may be carrying. The physical encumbrance of the gear is likely to restrict movement, both in terms of degrees of freedom as well as overall speed of movement. Accordingly, because of the potential influence of protective and other gear on the timing of activities (and possibly on the outcome of the engagement), we feel it necessary to account for the effects of encumbrances on the various phases of mission planning, rehearsal, and training. Methods for achieving this aspect of the battlefield simulation will involve computational techniques as well as strictly interface issues as discussed in Section 4.5.4. (Simply wearing the protective gear will not reflect the impediments to movement because, for the near term at least, the trainee will be seated at a workstation and not actually moving.)

The ability of the simulator to provide the trainee with a sense of whole body motion, and to allow the simulator to accept movements as information input, would provide a considerable degree of realism and "presence" that would be helpful in the latter stages of training where providing a feel for the operation is most important. Accordingly, we have rated body movement capabilities as desirable for MST, but unneeded for MPR and the early stages of CPT.

2.2.7 Other Interface Requirements

The category "other requirements" includes aspects of the simulation pertaining to the effects of physical stress that do not fall clearly into the above categories. This category includes fatigue (perhaps the most important factor), possible overheating if whole-body protective clothing is being worn, and overall physical degradation. Although these factors have not been emphasized in the documents reviewed in the course of this study, they merit serious consideration because of their potential influence on the course and outcome of the engagement.

To the extent that the behavior of a particular DI would be apparent to the other trainees participating in the simulation, the iconic representation of this DI should be appropriately modified to reflect movement restrictions (and possibly physical appearance). This modification is necessary whether the affected DI is a live trainee or strictly a computerized model. As

discussed is Section 3.2.9, computational as well as interface issues must be addressed.

Ideally, the (live) trainee would be made to physically feel the effects of the stress that would occur in the operational situation. In the absence of physical fidelity of this sort, the trainee should at least be given an indication of what stress(es) would be significant during the course of the mission, and the actions allowed by the simulation should be consistent with the performance degradation that would accompany such stress.

Specific recommendations concerning the application of VETT to meeting the various training requirements for CPT, MPR, and MST are given in Chapter 4.

3. Overview of Virtual Environment Technology

Some of the important features of virtual environment technologies relevant to DI training are reviewed here. Various "levels" of technology are defined, ranging from that which is either currently in use or required near-term for the Close Combat Tactical Trainer (CCTT) to technologies still in early stages of development. The general applicability of these technologies to CPT, MPR, and MST is suggested. Technology requirements to meet specific training needs are discussed more fully in Chapter 4, and specific devices and their current and projected states of development are presented in Chapter 5.

Although this chapter focuses primarily on the human/computer interface, advances in computational capability will need to keep pace with advances in interface technology if full benefits of the latter are to be achieved. Selected modeling issues are discussed along with interface issues in this chapter. In general, more computational "horsepower" will be required to meet the informational needs of the advanced interface technologies.

3.1 Working Definition of "Virtual Environment"

A simple dictionary definition of the term virtual environment is not available. There are, however, a number of features and capabilities that tend to differentiate virtual environment training (VE) systems from standard simulators. These features, described more fully in Durlach et al. (1992) include:

<u>Multimodal.</u> A VE training system employs a number of modalities to transfer information to, and receive information from, the trainee.

<u>Interactive/adaptive.</u> A VE training system tends to be human-centered, whereas a simulator is system-centered. In particular, the information displayed by a VE system depends on the state of the human as well as of the system, whereas the simulator outputs are based solely on the state of the system.

<u>Reconfigurable in Software.</u> The objective of a VE training system is to create as much of the human/system interface in software as is possible, and to rely as little as possible on physical mock-ups.

<u>Generate Unnatural and Unrealistic Situations.</u> Because of the reliance on software to create the simulated world and generate cuing mechanisms, a VE system can be easily configured to provide enhanced or even unrealistic cues

whenever such artifices are believed to enhance learning rates or otherwise serve training purposes.

As described in the popular press, VE systems are often associated with a sense of "presence," i.e., the sense of actually being immersed in the situation being simulated. Currently insufficient data exists to determine whether or not presence is a necessary ingredient of a virtual environment; hence we do nor consider it for the purpose of specifying ICS requirements. The need for research on the concept of presence is discussed in Chapter 6.

3.2 The Current DI Simulation

To the best of our knowledge, the SIMNET combat training system is the only operational battlefield simulation system to include the capability for representing the DI. The SIMNET DI feature has limited capability and is intended mainly to allow the tank, infantry fighting vehicle, and air crews being trained to coordinate their operations with those of the DI, rather than to specifically train DI personnel. Nevertheless, this initial implementation serves as a baseline from which to develop a combat training system that fully integrates the DI into the battle environment.

In the current implementation, a single icon is generated representing the soldier in a standing, kneeling, or prone position in the context of the battlefield. A workstation allows an operator to control, using joysticks, the movements and other actions (e.g., weapons firing) of the simulated infantryman. An array of CRTs simulates the DI's view of the scene, with the scene changing in an appropriate manner as the infantryman moves about the battlefield and/or changes his gaze direction. The icon represents an entire squad (i.e., it has the firepower of the squad), and the workstation operator plays the role of squad leader.

Other participants in the simulated battle who are within line of sight of the infantryman will observe the DI icon in its proper location and size relative to other elements of the scene.

3.3 Levels of VE Complexity and Sophistication

A range of technologies is available and/or under development for each of the modalities of information transfer provided by VE training systems. In this section we briefly review these technologies, and we assign ratings of from 1 to 3 to their levels of sophistication and complexity. In general, Level 1 indicates a technology that is currently available and is either currently used in the SIMNET combat training system or specified for the contemplated CCTT. Level 2 represents the next step in technological development, Level 3 the next step beyond

that, and so on. Computational and modeling issues related to VETT are addressed along with technologies relating directly to the soldier/simulator interface.

Levels of complexity for various types of technology are defined in Table 2. In this table, and in the discussion that follows, "displays" is used to signify transfer of information from the simulator to the trainee. "Sensing" means information or control inputs explicitly provided by the trainee to the simulator or otherwise sensed through electro-mechanical devices.

To some extent, the various levels of technology can be implemented independently. For example, any of the technologies for visual display can, in principle, be implemented concurrently with any of the technologies for auditory display. Some constraints also exist, however, and where they do exist, we have tried to assign levels that are consistent with these constraints. For example, the use of a head mounted display with uniform resolution (Level 2) requires accurate sensing of head position (also Level 2 technology). The next level of visual display complexity requires the next level of visual sensing capability.

A given technology level does not necessarily imply a given level of required developmental effort from one modality to the next. For example, providing the next step in haptic interfaces beyond what is currently available may require substantially greater effort than providing the next advancement in auditory interfaces. To associate time, cost, etc. with providing the various levels of technology, one must look at the specific technologies involved (see Chapter 5).

In this section we review the general characteristics of the various VE technologies directly applicable to training the DI. Specific devices are discussed in Chapter 5.

3.3.1 Visual Display

Visual displays may be categorized into head-worn displays in which the display surface moves as the wearer's head moves, and external ("off-head") displays that are fixed to the surrounding environment. Both types are considered here.

3.3.1.1 Level 1: Multi-Screen

As noted above, a multi-screen visual interface is provided for the DI operator in the current implementation of SIMNET. We consider this format as the Level 1 technology for the visual display.

In order to extend the effective field of view (FOV) of the display, the trainee (who sits at a workstation) has a joystick to slew the display to reflect changes in head orientation. This arrangement is perhaps suitable for simulating the view through a gunsight which has a relatively limited slew rate, but it is not feasible for reflecting rapid head rotation. An obvious drawback is that a hand that might otherwise be available for some other activity is needed to inform the simulator about the DI's head rotation.

In principle, additional CRTs could be used to extend the horizontal FOV to 360 degrees and to increase the vertical FOV to approximate the view provided by a dome-like display. This configuration would allow natural head movements, but at the cost of hardware and computational burden to provide a full 360 degree FOV.

A multi-screen or dome arrangement obviously cannot provide a stereoscopic view, since both eyes have the same view. Furthermore, without head position sensing (which is not considered Level 1 technology), the visual scene is independent of the position and orientation of the trainee. The trainee is thus deprived of parallax cues so useful for depth perception (i.e., the apparent shift in relative positions of objects as the trainee shifts head position). In addition, in order to provide a view corresponding to the trainee's body stance, the trainee must explicitly indicate body position (standing, kneeling, or prone) to the simulator via a mechanical control device.

Visual display technology at this level is probably adequate for activities that require viewing scenes at a distance where the human's fixation point is relatively fixed or slowly moving, such as viewing the scene ahead to detect landmarks or otherwise assesses the battle situation. As noted above, Level 1 display technology also appears adequate for simulating the view through the sight of a large weapon having a relatively limited slew rate.

3.3.1.2 Level 2: Helmet-Mounted Display (HMD), Low Resolution

A helmet-mounted display (HMD) provides the means for extending the FOV without continuously generating a full 360-degree visual scene. Ideally, the instantaneous FOV provided by the display spans the FOV of the unaided eye, and the visual scene displayed by the HMD reflects the trainee's instantaneous head position and orientation. This display format requires sensing of head position and orientation as discussed in the Section 3.3.2.1

Table 2

Levels of Virtual Environment Training Technology

Level 1	Level 2	Level 3
a) Visual Display		
Multi-screen	Helmet-Mounted Display, low resolution	Helmet-Mounted Display, high resolution
b) Visual Sensing		
(none)	Limb and body position	Level 2 + eye position
c) Auditory Display		
Battlefield sounds provided by speakers	Battlefield sounds provided by speakers and headphones	----
d) Auditory Sensing		
(none)	Limited speech recognition	Advanced speech recognition
e) Haptic Display		
(none)	Programmable specialized control devices	Programmable general-purpose control devices
f) Haptic Sensing		
Joystick standard control panel devices	Programmable specialized control devices	Programmable general-purpose devices
g) Whole-Body Movement		
(none)	Simulated large-volume movement through movement in place	Sensory stimulation involving no motion

Table 2 (Concluded)

Levels of Virtual Environment Training Technology

Level 1	Level 2	Level 3
h) Dismounted Infantry Models: Biomechanical Articulation		
Perspective view of icon appropriate to stance	Level 1 + low-fidelity articulation of head and limbs	Level 1 + fully animated icon
i) Dismounted Infantry Models: Influence of Physical Condition		
Indication of killed, wounded, or operational	Level 1 + movement speed degraded as appropriate	Level 2 + appearance modified as appro-priate
j) Physical Condition of Trainee		
none	Movement speed speed degraded as appropriate	Level 2 + artificially induced stress

By keying the generated scene to head position and orientation, the training system can provide visual parallax cues, and the trainee need not manipulate a mechanical device to reflect head movements. The HMD can also provide stereoscopic viewing by generating separate scenes for each eye. If the head-position sensing system operates over a sufficient spatial volume, the view can be synchronized with the trainee's actual body stance.

In addition to providing a generally more faithful cuing environment than the Level 1 multi-screen technology, the HMD is required for activities in which the soldier makes rapid head movements (e.g., rapid scanning movements right and left to determine the location and activities of other squad members or of nearby enemy soldiers). This technology is also required where the view depends critically on the position of the head relative to other objects (e.g., when the soldier momentarily exposes himself to view the scene then ducks behind cover).

We differentiate between Level 2 and Level 3 technology in terms of display capabilities (resolution, delay, update rate, etc.) that are achievable at reasonable cost. Level 3 is defined as sufficiently close to human perceptual capability as to not be a limiting factor for training purposes. Level 2 is defined as substantially less than Level 3 but still acceptable for most training proposes. (Clearly, there is a continuum of capabilities and the assignment of two levels is an artifice to help organize the discussion.)

In general, Level 2 or 3 technology is required for simulated battle conditions in which the DI is required to make rapid scanning movements over a large visual angle. Use of the HMD does, however, place significant additional requirements on simulator capability. To the extent that physical controls are used, the HMD must display images of these controls that correspond to and are in registration with the images that would be seen if the controls were viewed directly. If the controls are virtual, the HMD must still display an image of the control in space, the control must be operable, and the feel of the control must be provided by an appropriate haptic display mechanism (e.g., an instrumented glove or exoskeleton).

One caveat concerning the use of a HMD by a person sitting at a workstation or otherwise relatively fixed in space is the potential for "simulator sickness"; i.e., the feeling of discomfort or even nausea that often occurs when the body's vestibular motion sensors do not confirm the apparent motion reflected in the visual scene. This phenomenon is discussed later in Section 3.3.7. (See also Section III-A-1-e-ii of Durlach et al., 1992).

A comparison of the visual system specifications for the CCTT with specifications for various HMD's in development or prediction (Durlach et al., 1992) shows that the visual resolution for Level 2 technology is not as good as that provided by Level 1 (as exemplified by the CCTT). The resolution (more precisely, granularity) specified for the DI station at unit magnification is 6.7 arc-minutes between discernible lines, whereas relatively low-cost (under $20K) HMDs show resolutions in the range of 10-20 arc-minutes.

The high-end Level 3 devices (discussed below) show performance superior to both Level 1 and Level 2 technology, with actual or expected resolutions between 1 and 5 arc-minutes.

3.3.1.3 Level 3: Helmet-Mounted Display, High Resolution

The visual display system must provide detail with sufficient resolution to correspond to the human's view of the actual visual scene. In order to accommodate favorable viewing conditions, an HMD operating with uniform resolution must therefore provide a high resolution over the entire FOV. HMDs with limited resolution are available now, but the level of resolution available at a cost that would make it feasible to equip a team is probably not adequate to meet training requirements. Increasing the resolution also imposes a greater computational burden to respond to scene changes associated with head and body movements. One of the technical challenges is to be able to achieve satisfactory resolution and computational efficiency to meet training requirements.

At the cost of increased complexity in the soldier/simulator interface, one way to achieve computational efficiency and required resolution is to allow a non-uniform resolution that matches the spatial distribution of the resolving power of the eye. That is, relatively high resolution is provided only in the central viewing region, with much lower resolution used to fill the remaining FOV. This configuration requires that eye position (as well as head position) be sensed so that the high-resolution display area is synchronized with the trainee's look point.

Another way to increase display effectiveness and avoid the necessity to measure eye movements is to have differential resolution with the central high-resolution area sufficiently large to cover the portion of the FOV that the human will normally scan without making head movements. Finally, advances in computing power may allow Level 3 display technology to be achieved with high resolution in the entire FOV.

The ability of the HMD to present a high-resolution display is dependent, in part, on the ability of the simulator to provide information of necessary quantity and timeliness. (That is, there is no point in increasing the pixel resolution if the

pixels cannot be updated accurately and rapidly enough to provide a smooth visual scene, with minimal delay).

First, a factor of two increase in information rate is required to provide separate images to the two eyes to allow stereoscopic vision. Beyond that, the detail contained in the terrain models, and the rate at which the displayed information can be updated, must be increased to support the conditions for which Level 2 and 3 displays are recommended. For example, detailed models for buildings and building interiors will be needed for simulations of urban fighting.

Applicability of this technology is the same as that described for Level 2.

3.3.2 Visual Sensing

Measurements of position and movement of various body parts, especially the head, are required to implement Level 2 and 3 display technologies. The mechanisms for performing the sensing reviewed in this section are largely electro-optical and electro-magnetic, where we use the term "electro-optical" to include analysis of light patterns provided by light emitters attached to helmets, gloves, etc. Although it is an active area of research, true "visual sensing" of body position and movement in terms of computer analysis of complex visual scenes is not being considered for near-term payoff in VETT.

3.3.2.1 Level 1: None

Sensing of this category is not currently in use for DI simulations and is not required for Level 1 visual display technology. We therefore classify the least complex technology as Level 2, which, as we show below, is consistent with our classification of visual display technologies.

3.3.2.2 Level 2: Limb and Body Position

Head movement must be measured in order to implement the Level 2 display technology (HMD with uniform resolution). That is, it is imperative to know the trainee's point of regard in order to display the appropriate visual scene. Training activities that require Level 2 visual display capabilities therefore also require level 2 sensing capability.

"Visual" technologies for sensing finger, hand, arm, and body position are essentially the same as for measuring head position and are all considered part of the Level 2 visual sensing technology. This capability, coupled with Level 3 haptic display capability (Section 3.3.5), allows the generation of virtual controls where computer-generated imagery is used to display the control device and a general-purpose mechanical

device is used to provide the appropriate physical feel and operation.

This capability would allow the trainee to behave in a more operationally-realistic manner. For example, instead of pressing a button to indicate their stance to the computer, trainees could actually change stance. Similarly, actual hand and arm signals could be used instead of mechanically issuing commands to (computerized) troops.

Haptic sensing provides an alternative to sensing limb and body position as discussed in Section 3.3.6.

3.3.2.3 Level 3: L2 + Eye Position and Movement

The next level of technological complexity for electo-optical and electromagnetic sensing involves adding the capability to measure eye movement and fixation point. This level is required to implement the Level 3 visual display capability, which is to use a HMD with a resolution that varies with angular distance from fixation point.

3.3.3 Auditory Display

The simulator will provide radio and telephone voice communications through headsets and/or telephones. Since this technology is currently in place and is not considered in need of significant advancement, we do not consider it further. With regard to VET, our discussion of auditory displays relates to the presentation of battlefield sounds and, in general, acoustic inputs other than voice through radio/telephone.

3.3.3.1 Level 1: Battlefield Sounds Presented by an Array of Speakers

Because the CCTT specification calls for battlefield sounds to be provided specifically by speakers, and because the ability to identify and localize auditory events is considered necessary for training the DI, speaker presentation of battlefield sounds is defined here as the Level 1 technology. This technology requires an array of speakers of sufficiently high fidelity to accommodate the range of sounds relevant to the battle environment (including low frequency vibrations caused by vehicle movement, weapons operation, explosions, etc.). Direct voice communication between trainees might also be provided in this manner.

Because the sounds generated by the simulator are independent of the location of the listener, use of this technology requires that the physical (not simulated) movements of the trainee be relatively restricted (e.g., the trainee remains at a workstation). As is the case of Level 1 visual

display technology, Level 1 auditory display technology is best adapted to replicating sounds that are located relatively far from the trainee's simulated location in the battle environment. As we discuss below, Level 1 is also restricted to situations in which accurate localization in the vertical dimension is not required.

Advanced auditory displays will require additional computer modeling for support, even for Level 1 display technology. The current implementation of SIMNET has only limited acoustic modeling. Only a few weapons are modeled acoustically, and for weapons other than those operated by the trainee or their team, the sound of the shell landing is replicated (if it lands nearby), but not the sound of the weapon being fired. Some sounds are amplitude-adjusted to account for the range of the sound source. No directional information is provided.

Given the ability of a speaker array to provide reasonable acoustic fidelity (at least in two dimensions), models will be needed for more battlefield sound sources, and the speaker array will have to be driven in a manner that allows DIs to localize the sounds (as well as they would be able to localize them operationally).

3.3.3.2 Level 2: Battlefield Sounds Presented by Speakers and Headphones

Speaker arrays are limited in a number of respects with regard to providing adequate representation of the battlefield auditory environment:

Although most of the battlefield sounds are at or near ground level, the simulation of close air support provides above-ground sound sources that the DI may wish to localize. Adding the vertical dimension to a speaker array adds considerably to the amount of equipment needed.

The speakers will be located a number of feet from the trainee, which may limit the accuracy with which sounds close to the DI can be represented.

Difficulties in controlling the acoustic environment (e.g., wall reflections, extraneous sounds) may significantly degrade the simulated sound environment.

A substantial amount of space will be needed for each DI station to accommodate the speaker array, and substantial shielding will be needed to prevent acoustic interference with nearby DI stations.

Level 2 technology (speakers plus headphones) is therefore recommended for the presentation of battlefield sounds when the

sounds are relatively close to the DI (e.g. a companion is talking from a short distance away), or when localization in the vertical dimension is important. We recommend a combination of speakers and headphones - rather than headphones alone - to provide a practical mechanism of providing some of the very low-frequency sounds and vibrations representative of battle that would be felt rather than heard.

Because very low-frequency sounds are not used for localization, a reasonable scheme is to present low-frequency sounds through speakers (perhaps a single speaker will suffice) and use headphones for presenting sounds at higher frequencies. The headphones would also be used for ratio/telephone communication.

Use of Level 2 technology for auditory display will require the concurrent use of Level 2 sensing technology to track the trainee's head location and orientation.

3.3.4 Auditory Sensing

For this discussion, "audio sensing" refers to computer interpretation of auditory inputs from the trainee, not the ordinary use of a microphone to allow one individual to talk to another.

3.3.4.1 Level 1: None

Because speech recognition is not in use for the current DI simulation and is not required for the CCTT, the least complex auditory sensing technology is rated Levei 2.

3.3.4.2 Level 2: Limited Speech Recognition

If the technology were readily available, trainees in the role of squad leader would command their computerized DI "troops" in the same manner as they would in the field, including by voice when appropriate. In order to facilitate this mode of operation, the simulator must be able to interpret spoken commands for DI modules that are strictly computer modeled. (Microphones and headsets would be used to allow one trainee to communicate with another.)

Level 2 auditory sensing consists of speech recognition capability that is easiest to implement - specifically, systems that are speaker dependent and have a relatively small vocabulary. Such systems would require "training" for individual speakers and would require the use of a limited vocabulary. Because of the small number of trainees using a particular simulator host at a given time, and because of the specialized nature of the training, these restrictions do not seem particularly burdensome.

While not considered necessary for the training purposes considered in this report, the ability to command the computerized squad members via speech might prove useful for mission-specific training where it is more important for the trainee to have the feel of the situation. Should computerized voice recognition become an operational capability for military operations (e.g., the squad leader issues voice commands to a piece of equipment), this capability would need to be included in the simulation.

3.3.4.3 Level 3: Advanced Speech Recognition

The next level of technological development removes some of the restrictions associated with Level 2. The Level 3 system, for example, is defined as being speaker independent and able to handle a larger vocabulary. Level 3 technology is applicable to the same situations as Level 2 technology. Presumably, Level 3 would be preferred over Level 2 if the technologies were equally available and of comparable cost.

3.3.5 Haptic Display

"Haptic display" refers to information provided by the ICS to the soldier through the senses of touch and kinesthetics, where this information is based on the states of the simulated world and of the trainee.

3.3.5.1 Level 1: None

Although standard control devices such as switch panels, touch screens, joystick, etc. provide a form of physical feedback to the human, this feedback is not dependent on the state of the simulation. Such devices are therefore not considered to be "haptic displays" as we have defined the term. Situation-dependent haptic feedback is not currently in use for DI simulations and does not appear to be contemplated for near-term simulators. Consequently, we classify the least complex haptic display technology as Level 2.

3.5.2 Level 2: Programmable Specialized Control Devices

Level 2 technology includes specialized control devices that provide some form of state-dependent feedback to the trainee. An example of this technology is the artificial force-feel system provided in aircraft simulators (and, in some cases, in the operational aircraft) in which the force/displacement relationship of the control stick or wheel is modified to reflect current operating conditions such as airspeed and angle of attack.

Level 2 technology is potentially useful for indicating to trainees either their own degraded mobility or the degraded

mobility of the simulated troops under their command due to the effects of encumbrances or other stresses. For example, if the displacement of a programmable joystick is used to command the speed of movement for the simulated DI, the force/displacement relationship could be adjusted to increase the amount of force required to generate a given movement speed, thereby providing a physical indication of the degraded state of the DI. (Additional indications of degraded capability would typically be provided computationally and visually through a reduced maximum speed.)

Level 2 technology is considered unnecessary for planning purposes and strictly procedural training, but it is considered desirable for mission-specific training where it is more important for the trainees to have a physical feel of the situation and to be more acutely aware of their physical limitations.

3.3.5.3 Level 3: Advanced General-Purpose Devices

Level 3 technology, which is more in the spirit of "virtual environments" as defined above, employs a general purpose device for providing haptic feedback. That is, the hardware is not a replica or model of a specific control device, but is a relatively general-purpose device that can be programmed to represent a number of control devices or other sources of physical contact with the soldier.

Level 3 devices may be classified into two broad groupings: "ungrounded" (more precisely, grounded to the human) and "grounded" (i.e., grounded to the external environment). Ungrounded devices are worn by the human and are best used to represent self-equilibrating forces (e.g., squeeze a trigger); a grounded system is recommended when the forces are unbalanced (e.g., push against an object fixed in space, bump into a wall). Specific devices are reviewed in Durlach et al. (1992).

Level 3 technology is, in general, applicable wherever Level 2 technology is considered applicable. Depending on the state of development, a combination of Level 2 and 3 technologies might be appropriate (e.g., a programmable joystick plus an ungrounded exoskeltal device).

The use of Level 3 haptic sensing technology will require the use of Level 2 or 3 visual display technology. Because the general-purpose device will not physically resemble the object whose characteristics it is simulating and may not be locatable at a corresponding location in space, a HMD will be needed to display the simulated objected at its proper location and with its proper appearance.

3.3.6 Haptic Sensing

Three levels of technology are considered for haptic sensing: (1) standard control devices, (2) programmable specialized control devices, and (3) programmable general-purpose devices.

3.3.6.1 Level 1: Standard Control Devices

Until such time as speech recognition systems are sufficiently reliable and cost-effective to be used as the mechanism for accepting commands from the trainee, mechanical input devices will be needed to allow the squad leader to command simulated troops, to control the operation of weapons and other equipment, and to provide commands for manipulating objects on the battlefield.

We define Level 1 technology to be that which is in use currently and expected to be in use over the near future; namely, standard control devices such as (non-programmable) joysticks, switch panels, and touch screens, even though such technology does not literally fall within our working definition of VET.

3.3.6.2 Level 2: Programmable Specialized Control Devices

The Level 2 haptic display devices, by their very nature, serve as both input and output devices. Therefore, we classify such devices as Level 2 sensing technology as well.

3.3.6.3 Level 3: Programmable General-Purpose Devices

Level 3 haptic sensing devices have similar characteristics to Level 3 haptic display devices: (1) they are programmable, (2) they do not in general physically resemble the objects or features being simulated, and (3) they may be classified as either ungrounded or grounded. Level 3 devices fall in the category of "virtual controls".

As with the Level 2 devices, some Level 3 devices may serve both as haptic sensors and displays. For example, a set of linkages attached to the arm could serve both as a means for the simulator to detect arm position and movement, as well as a means for providing programmable resistance to movement to simulate a particular encumbrance. But some devices may serve strictly as sensors (e.g., a mechanical system for sensing head position for use with a HMD), and others may be display devices only (e.g., vibrators embedded in a glove to provide various touch sensations).

Level 3 sensing technology is generally applicable to the same training environments as Level 3 display technology (except

that use of a HMD does not necessarily require the use of Level 3 haptic sensing technology).

3.3.7 Whole-Body Movement

There are two aspects to the incorporation of whole-body movement into the ICS: (1) the capability to provide the trainee with the sensation of movement, and (2) the use of body movement as an information input to the simulator. As discussed in Section 5.2.6, the trainee may achieve the sense of motion artificially through various forms of mechanical or electrical stimulation, or through a provision to allow substantial movement on the part of the trainee (e.g., a treadmill).

Note that whole-body movement here is conceptually different from that commonly associated with aircraft simulation. A moving-base aircraft simulator provides aircraft motion in up to six degrees of freedom, but the pilot is strapped to the seat. Thus, the pilot has relatively limited degrees of freedom for movement. For the DI, however, the platform is stationary, and any whole-body motion is that of the (dismounted) soldier with respect to the battlefield.

Three levels of technology are briefly considered below: (1) no representation of whole-body motion, (2) simulated large-volume movement through movement-in-place, and (3) sensory stimulation involving no motion. Actual whole-body motion in which the trainee undergoes significant translational movement is not considered practical for ICS simulation because of the large area that will be covered in normal military operations. Potential methods for handling whole-body motion are reviewed in Section 5.2.6.

The need for providing the sensation of whole-body motion will depend, in part, on the degree to which the lack of whole-body motion results in what is termed "simulator sickness". This phenomenon has been observed in fixed-based airplane and automobile simulators and is often ascribed to "cue mismatch" - i.e., a situation in which motion cues (or lack of same) provided by the vestibular sensory apparatus are inconsistent with the appearance of motion provided by visual cues. This phenomenon is believed to be exacerbated by especially compelling visual scenes.

The presence of simulator sickness may compromise training effectiveness in one of two ways: some trainees who are susceptible to this effect may not be able to train in the ICS, whereas others may learn to overcome the effects by adopting response behaviors that are inappropriate to the operational situation, thereby reducing positive transfer of training.

Not enough is known at present to predict the occurrence and degree of simulator sickness that might result in a particular ICS training environment. For now, we shall base our recommendations on the assumption that sickness is not a significant issue. Should experience prove this assumption false, then our recommendations will have to be modified to require a tight coupling between visual display technology and whole-body motion cuing for those training scenarios where a significant motion/visual cuing mismatch would occur (i.e., provide realistic whole-body motion cuing whenever a high-fidelity visual display is used, or, if the motion-cuing technology is not available, then avoid using a high fidelity visual display.

3.3.7.1 Level 1: None

Provision for whole-body motion for the DI in a combat simulation is not envisioned in the near future, nor is it deemed to be required for the types of training considered in this report (assuming simulator sickness is not an issue).

3.3.7.2 Level 2: Simulated Large-Volume Movement

Level 2 technology provides the trainees with the capability to engage in physical activity appropriate to movement in space (e.g., walk or run on a treadmill) without significantly changing the location of their body in real space. This capability would be most appropriate for MST (and perhaps advanced stages of CPT) where the sensation of motion would most likely benefit the training process.

3.3.7.3 Level 3: Sensory Stimulation Involving no Movement

Level 3 technology provides the sensation of motion artificially through stimulation of one or more senses in a way that involves no actual motion. One such scheme involves vibration of appropriate muscle groups. (See Section III-A-1-e-i of Durlach et al., 1992) Applicability of Level 3 technology is the same as for Level 2.

3.3.8 DI Models: Biomechanical Articulation

The question of how to implement the iconic representation of the DI is a modeling issue rather than a purely interface issue; that is, any of the visual display technologies discussed above are presumed capable of presenting the DI icon in any of the levels of detail considered here.

Three levels of modeling complexity are considered. Level 1 (currently implemented in the SIMNET training system) provides a perspective view of an otherwise fixed representation of the DI in one of three stances: standing, kneeling, and prone. That is,

the view shown to the trainee changes with the relative position of the simulated DI, but, for a given stance, the same computerized model is used. Level 2 technology adds to Level 1 by adding articulation of the head and limbs to provide the observer a better indication of the DI's current activity. We expect that initial implementation of an articulated DI will exhibit somewhat cartoonish behavior in its simulated movements. Level 3 technology envisions fully animated DI models having smooth realistic movements.

As discussed in Chapter 4, Level 2 and 3 technology is recommended primarily for battle situations in which intense activity is conducted close by, but only if a DI icon corresponds to a single soldier.

3.3.9 DI Models: Influence of Physical Condition

Reorganization after losses is an important part of the warfare. At the very least, the ICS must indicate to squad leaders the gross operational status of their troops (killed, wounded, fully operational). In addition to casualties, performance degradations suffered by operational troops will influence the conduct of operations and may play a key role in the outcome of the battle. Again, this is largely a modelling issue.

3.3.9.1 Level 1: Gross Indication of Operational Status

At a minimum, the ICS must indicate current troop strength. The way in which this is handled will depend on whether the DI icon represents a single soldier or a group of soldiers. If each soldier is individually represented, color (and/or shading) and positional coding of the icon can be used to indicate whether the DI is operational, wounded, or killed. As casualties are removed from the battlefield, the corresponding icons are removed from the display.

Various forms of coding (e.g., color, shading, size, shape) may be used to indicate the strength of a DI unit (e.g., fire team, squad). Troop strength below the minimum necessary for effective combat might be indicated either by removal of the corresponding icon, or, if the possibility of integrating the remaining soldiers with other units exists, by coding the appearance of the icon and eliminating its firepower until the required reorganization has taken place.

Level 1 technology is recommended for CPT, which is expected to be focussed on a particular task rather than on a full mission scenario. Fatigue and depletion of human consumables (e.g., food and water) is not likely to be an issue for this type of training exercise.

3.3.9.2 Level 2: L1 + Degraded Movement Speed

Troops that are not casualties in terms of being wounded or killed may nevertheless suffer loss of performance due to various sources of physical stress (e.g., fatigue, depletion of consumables such as food and water). The calculation of the degree of stress and its impact on performance is, of course, a modeling and computational issue. To some extent, the representation of these effects to the observer is also a modeling issue (e.g., reduction in maximum movement speed and marksmanship). As suggested in Section 3.3.5.2, this reduced capacity might be indicated in part through the characteristics of control input that the trainee uses for commanding troops.

Level 2 technology is recommended for planning and mission-specific training, where it is important to account for all significant factors that influence the speed and outcomes of operations.

3.3.9.3 Level 3: L2 + Appearance Modified as Appropriate

The squad leader will often have some indication of the effects of fatigue and other stress on troops other than by a reduced ability to perform tasks; specifically, through verbal communication and through the appearance of the soldier. Although supporting a conversation between live trainees and simulated Dis does not appear feasible within the near future, modification of the appearance of the DI icon is well within current capabilities. We therefore suggest a Level 3 technology that visually indicates the degraded state of the DI through, say, color coding and/or shading. Level 3 technology is applicable wherever Level 2 technology is considered applicable.

3.3.10 Physical Condition of the Trainee

In addition to being made aware of the effects of stress on the simulated DI, trainees must be made aware of the same effects on themselves (i.e., the discomfort and/or loss of performance that they would encounter in the actual combat situation). This consideration applies mainly to trainees who are individually represented in the ICS simulation. (In the current SIMNET configuration, and in the initial implementation of the CCTT, the squad leader is considered one member of an operational unit represented by a single icon, thus precluding separate treatment of the trainee playing the role of squad leader.)

There is no meaningful Level 1 technology. The soldier represented by the trainee is either fully operational, in which case no special action is taken, or non-operational (killed or wounded), in which case the trainee's display can be turned off and their controls rendered ineffective until another (or the same) trainee assumes the role of a simulated replacement.

Level 2 technology is similar to that suggested for the simulated DI; namely, that the (simulated) reduced capability of the trainee be reflected through the physical characteristics of the control device that is used to simulate movement over the battlefield, and that maximum achievable speed be reduced.

Level 3 technology for the (live) trainee is different from (although parallel to) Level 3 technology for the DI because the task is now to make the trainee directly feel the effects of stress. Because trainees will, for the foreseeable future, interact with the ICS through workstations at which they are seated, the simulated activities are not likely to induce the same degree of fatigue, etc., in the trainees as would occur if they were actually participating in activities of the type being simulated. Some form of substitute stress inducement is suggested, such as heating the trainee's enclosure or clothing. To our knowledge, this aspect of VETT has not been pursued.

4. Application of Virtual Environment Technology to Training

In this chapter we discuss in greater detail the ICS training requirements summarized in Chapter 2, and we suggest the appropriate level of VE technology to be applied in meeting each training requirement. The reader is referred to Chapter 3 for definitions of the levels for the various training technologies. These levels have been summarized in Table 2.

Table 3 summarizes our recommendations concerning the nature of the VE technology required to meet the various training needs. (Specific devices are discussed in Chapter 5.) The organization of this table largely parallels that of Table 2 in that it retains the same major technology-oriented headings, but the entries in terms of training requirements are organized similarly to those of Table 1.

To obtain a complete picture of the VE technology needed to meet a specific training need, the reader first uses Table 1 to assess the importance of providing the relevant simulator capability. If this capability is considered either desirable or necessary, Table 3 is then used to suggest the type of VE technology to be applied. (Models for the DI elements - not mentioned in Table 1 - are necessary for all training purposes.)

A "-" indicates that a requirement does not exist and that a technology definition is thus not relevant. A "+" indicates that a higher level of technology than indicated is either generally desirable or necessary under certain conditions. A "*" directs the reader to the text for a discussion of technology requirements that do not readily conform to those listed in Table 2.

The technology levels shown in Table 3 are considered the minimum adequate for training purposes. In general, a higher level includes the capabilities provided by lower levels; thus, training needs should not be compromised (and may often be enhanced) by using a higher level than specified in the table.

4.1 Visual Display

In general, Level 2 and 3 technology (head-mounted display) is required for simulated battle conditions in which the DI is required to make rapid scanning movements over a large visual angle. For other situations - which account for most of the visual perceptual activities - Level 1 (multi-screen) should be adequate to meet training needs.

Because Level 3 technology (high-resolution HMD) can be applied wherever Level 2 (low-resolution HMD) is called for, all recommendations for Level 2 visual display technology are appended with the plus sign.

Table 3

Summary of VE Technology Requirements for ICS Interface

Training Requirement	**Training Purpose and Level of VE Technology Required**		
	CPT	**MPR**	**MST**
a) Visual Display (Level 1 = Multi-screen display) (Level 2 = HMD, Low resolution)			
Perform navigation.	1	1	1
Assess physical environment with respect to cover and concealment.	1+	1+	1+
Assess suitability of terrain and other aspects of physical environment for troop & vehicle movement.	1	1	1
Assess likelihood of enemy travel paths.	1	1	1
Conduct operations in an urban environment.	2+	2+	2+
Conduct non-urban close-in operations.	2+	2+	2+
Determine status of friendly troops.	1+	1	1+
Determine status of enemy troops.	1	1	1
Distinguish status of friendly and enemy vehicles.	1	1	1
Distinguish/Identify friendly and enemy forces.	1	1	1
Assess status and effects of weapons.	1	1	1
Recognize and assess status of various battle-field details (barbed wire, obstacles, etc.)	1	1	1+
Read navigation aids.	1	1	1

\+ A higher level may be desired or needed. See text.

Table 3 (Continued)

Summary of VE Technology Requirements for ICS Interface

Training Requirement	Training Purpose and Level of VE Technology Required		
	CPT	MPR	MST
b) Visual Sensing (Level 2 = Limb and body position)			
Conduct operations in an urban environment.	2+	2+	2+
Conduct non-urban close-in operations.	2+	2+	2+
Command troops via hand and arm signals.	-	-	2+
c) Auditory Display (Level 1 = None) (Level 2 = Battlefield sounds provided by speakers and headphones)			
Communication among trainees.	1	1	2
Conduct operations in an urban environment.	1+	1+	2
Conduct non-urban close-in operations.	1+	1+	2
Determine status of friendly & enemy vehicles.	1	1	1
Assess status and effects of weapons.	1	1	1
Be subjected to disruptive effects of battle-field noise.	1	1	1
d) Auditory Sensing (Level 2 = Limited speech recognition)			
Command simulated troops.	-	-	2+

+ A higher level may be desired or needed. See text.
- Not applicable.
* See text.

Table 3 (Continued)

Summary of VE Technology Requirements for ICS Interface

Training Requirement	**Training Purpose and Level of VE Technology Required**		
	CPT	**MPR**	**MST**
e) Haptic Display (Level 2 = Programmable specialize control devices) (Level 3 = Programmable general-purpose devices)			
Receive feedback concerning physical condition of troops.	-	-	2
Receive realistic feedback from control & manipulation of weapons, equipment, & objects.	3	-	3
Assess battlefield conditions.	3	-	3
Be subjected to physical and cognitive limitations of protective & other gear.	*	*	*
f) Haptic Sensing (Level 1 = Joystick & standard control panel devices)			
Command simulated troops.	1	1	1
Perform control & manipulation of weapons, other equipment, and objects.	1+	1+	1+
g) Whole-Body Movement (Level 3 = Sensory stimulation involving no motion)			
Perceive a sense of body movement.	-	-	3
h) DI Models: Biomechanical Articulation (Level 1 = Perspective view of icon appropriate to stance)			
Conduct operations in an urban environment.	1+	1+	1+
Conduct non-urban close-in operations.	1+	1+	1+

\+ A higher level may be desired or needed. See text.
\- Not applicable.
* See text.

Table 3 (Concluded)

Summary of VE Technology Requirements for ICS Interface

Training Requirement	**Training Purpose and Level of VE Technology Required**		
	CPT	**MPR**	**MST**
i) DI Models: Influence of Physical Condition (Level 1 = Indication of killed, wounded, or operational) (Level 2 = Level 1 + movement speed degraded as appropriate)			
Perceive effects of degraded physical condition of simulated DI.	1	2+	2+
j) Physical Condition of Trainee (Level 2 = Movement speed degraded as appropriate)			
Suffer physical and cognitive limitations on performance due to degraded physical condition.	_	2+	2+

+ A higher level may be desired or needed.
- Not applicable.
* See text.

We review below some of the procedural skills that need to be trained, and we identify the associated VE display technology requirements. As we noted earlier, to accommodate an adequate range of combat activities and environments, the ICS should be capable of presenting views as they would be seen from the standing, kneeling, or prone positions using unaided viewing, binocular viewing, or night vision equipment as appropriate. (Issues relating to the scene content and point of view concern mainly computer modeling and image generation capabilities rather than the soldier/simulator interface hardware.)

4.1.1 Perform Navigation

The DI will use both natural and cultural features and landmarks for navigational purposes. (Use of maps and other documents is discussed later.) Natural features include mountains, hills, valleys, bodies of water, fords, culverts, underpasses, etc.; cultural features include roads, buildings, bridges,etc. For strictly navigational purposes, general recognition of the specific object is sufficient (e.g., one particular lake should be differentiable from another particular lake); a high level of detail is often unnecessary. Because navigational activities are usually relatively slow-moving and are based on visual scenes at some distance, Level 1 technology is considered sufficient.

4.1.2 Assessment of Cover and Concealment

The ability to obtain and use cover and concealment is an especially important part of warfighting. To a large extent, offensive ground action consists of movement from one covered position to another. There are two aspects of cover and concealment that must be considered when designing the battlefield simulator: (1) the ability of the DI leader to assess the potential for cover and concealment when planning tactical activities, and (2) the presentation of the DI's view of the battle area as affected by those elements of the battlefield that provide cover and concealment. Cover and concealment can be provided by both natural terrain features (hills, boulders, vegetation, etc.) as well as by cultural features such as buildings and other large structures.

To facilitate tactical planning, the visual interface must allow the squad or platoon leaders to make a realistic assessment of the nature and location of cover and concealment on those portions of the battlefield that lie between their current position and their desired position(s). Leaders must also be able to estimate the travel time between successive positions, which means that they must be able to estimate the time required to move troops (and possibly vehicles) over the terrain. Because planning activity of this type is based on visual cues that are at some distance ahead (and thus do not require rapid scanning),

Level 1 technology is considered sufficient for training the DI in the assessment of cover and concealment.

The use of concealment while carrying out an operation is a separate consideration and requires simulator interaction at the level of the individual soldier. Here the soldier/simulator interface must also correctly present the visual scene to the DI as it would be influenced by those aspects of the terrain pertaining to cover and concealment, taking account of the location and body position of the DI. For example, DIs who crouch behind a rock to acquire cover may not be able to see the enemy.

In general, full physical fidelity of the visual simulation will not be required for cover and concealment. Rather, the object should be portrayed with sufficient fidelity so that the trainee knows what it is and so that its influence on the trainee's ability to see and be seen (and be protected from fire) is properly depicted. For example, if the trainee is undergoing CPT and needs to use a building for cover, the building should be recognized as such. If the training is related to a specific mission where the trainee is to use a specific building for cover, the displayed image should allow recognition of that specific building (to the extent that it would be recognizable given the anticipated visual conditions).

The level of visual display technology needed to train the DI in the use of cover and concealment (more generally, to provide a visual scene influenced by cover and concealment) depends on the details of the simulated engagement. If the troops are moving in a generally forward motion from one covered position to another, and the movements in and out of cover are relatively smooth, Level 1 visual display technology should be adequate. As discussed below in Section 4.1.4, Level 2 or 3 is needed for situations in which the DI darts in and out of cover or is otherwise presented with a rapidly shifting visual scene caused by self movement.

4.1.3 Assess Potential Travel Paths

One of the tasks of the DI decisionmaker is to assess the suitability of potential travel paths for friendly units and to assess the likelihood of enemy travel paths. To make this assessment, one must be able to determine the presence and condition of certain cultural (man-made) features such as roads and bridges. In the absence of such cultural features, the decisionmaker will have to assess the suitability of natural terrain for travel, including the travel of mechanized forces if required by the battle plan. The discussion here concerns orderly movements over relatively long distances where combat is not expected, as opposed to dashes from cover-to-cover during an assault or retreat under fire.

A number of details must be considered when planning a travel path for friendly forces. Many of the same judgements will be needed to determine where the enemy is likely to travel. Such details include topography, surface conditions, natural and man-made obstacles, and the potential for ambush (as determined in part by the cover and concealment available to the enemy).

The task of assessing potential travel paths for either friendly or enemy forces is mostly cognitive; only low-level physical activity is associated with obtaining the necessary visual cues. Level 1 technology is therefore considered sufficient for this aspect of training.

4.1.4 Conduct Operations in an Urban Environment

By "operations in an urban environment" we mean active military engagement (i.e., "street fighting") in which combat is conducted in and around buildings. This type of battle environment requires the soldier to make rapid movements (e.g., duck in and out of cover) and to make large and rapid scans (e.g., enter a room and scan for enemy soldiers), all on a relatively frequent basis.

This operational environment poses significant technical challenges for the ICS in terms of providing a suitable visual display environment that is compatible with having the trainee seated at a workstation. Given the restriction that whole-body movements are not accommodated by the ICS (a restriction likely to be in force over the near term), there are basically two approaches for providing control and display information: (1) Level 1 technology in which all DI movements are indicated through a joystick and the visual display consists of multiple screens, and (2) combined Level 1 and Level 2 technology in which DI movements are indicated both through a joystick and sensed head motions and a HMD displays the visual scene.

Let us review first the use of only Level 1 technology. In this mode, the trainee typically uses one or two joysticks to indicate translational movements and head point of regard. In order to accommodate long-range movement over the battlefield, the translational joystick must be interfaced in a manner that allows the trainee to indicate the direction and rate of motion, with the result that this control input is not convenient for indicating short and rapid motion (i.e., momentarily peek around a building). Furthermore, the rapidly shifting visual scene that accompanies this simulated head translation is likely to produce an interval of strobe-like images because of the finite update rate of the CIG. (Increasing the update rate would not solve the problem; it would provide a blurred image.) This visual artifact may well be disconcerting to the trainee (even to the point of inducing "simulator sickness") and may also interfere with the learning process. Blanking of the display during simulated head

motion would eliminate the strobe effects but would also be disconcerting to the subject and of negative training value.

Although the joystick (or joystick axis) that is used to indicate head rotation does not have the calibration problems associated with the translational mode, the strobing of a rapidly shifting image remains a factor. Use of a wrap-around array (or dome) avoids the need for the trainee to indicate head position to the simulator (and thus avoids the strobed image), but at the cost of the substantial computational burden to provide a continuous 360-degree presentation of highly detailed imagery. In addition, the wrap-around visual display will not provide parallax cues, because the computer has no mechanism for tailoring the visual scene to the subject's instantaneous head position without the associated control input.

Use of a HMD has the potential for providing parallax cues, avoiding the need for continuous 360-degree image generation, and reducing the perception of strobe effects associated with a rapidly shifting image. (The strobe effects may be present, but visual perception is markedly reduced during intervals of rapid eye movement.) Provided the required head (and possibly) eye tracking apparatus is implemented, the simulator "knows" the trainee's point of regard and can generate the appropriate visual scene.

Given that the trainee is seated at a stationary workstation, the joystick will still be needed in part to indicate the trainee's rotational movements. To illustrate the point, consider the situation in which the trainee travels north alongside a building, then turns to move west along the northern wall. If point of regard is indicated only by head rotation, the trainee must maintain a fixed head position of 90 degrees to the left. (The situation becomes especially acute when the trainee next decides to move south!) The solution is to use the joystick(s) to indicate whole-body translation and rotation, and the head sensing apparatus to indicate head movements.

4.1.5 Conduct Non-Urban Close-in Operations

The are numerous non-urban fighting environments that provide the type of action and close-in visual scenes that require Level 2 or 3 display technology. Examples include (1) darting in and out of cover or concealment, (2) attending in rapid sequence to squad members located on either side, and (3) close combat in which enemy troops may be located in any direction. For the reasons stated above, Level 2 and 3 display technology is recommended for this situation.

4.1.6 Determine the Status of Friendly Troops

When on maneuvers, DI leaders must assess the status of their troops - their location, movements, activities, and condition (i.e., killed, wounded, or operational). For troops within line of sight this assessment will be done visually. Level 1 technology is considered adequate for planning purposes and for near-term implementation of the CCTT, where a single icon will be used to represent an entire fire team (half of a squad). As the ICS is expanded to allow individual icons for each simulated DI, the HMD may be required because the friendly troops may be located on either side of the trainee.

4.1.7 Determine the Status of Enemy Troops

We refer here to the visual assessment of enemy troops located in their assembly area. The task is to assess their relative strength, their movements as a unit, and so on. Since this task usually requires a relatively narrow field of view located ahead of the trainee (or at least does not require rapid scanning movements), Level 1 technology is deemed adequate.

For enemy located close by, such as enemy fighting in close combat or POWs at hand, the considerations are the same as for assessing the status of friendly troops.

4.1.8 Determine the Status of Friendly and Enemy Vehicles

The DI must have the capability to determine the status of both friendly and enemy military vehicles; specially, the location and movement of vehicles and the extent to which the vehicles are operable. For friendly vehicles, the simulator must provide visual cues appropriate to whether the DI is mounted or dismounted, and it must provide details of the visual scene relevant to the act of mounting/dismounting the vehicle (e.g., ramps and doors).

The considerations imposed by the requirement to assess vehicle status relate to computational and image generation capability and not to VE hardware technology. For the most part, this type of cognitive activity should not require rapid physical activity, and Level 1 technology is deemed adequate.

4.1.9 Distinguish/Identify Friendly and Enemy Forces

The DI must be able to distinguish between friendly and enemy troops and vehicles - a task often difficult to perform in actual combat. The DI must pinpoint specific friendly vehicles (e.g., the specific vehicle to mount or get supplies from), and may also want to identify or classify enemy vehicles in view in order to assess capability and likely course of action. Again, these requirements imply certain computational capabilities

rather than interface capabilities, and Level 1 technology is considered appropriate for meeting this requirement.

4.1.10 Assess Status and Effects of Weapons

There are potentially a number of assessments the DI must make with respect to the condition and activity of friendly and enemy weapons. These include:

Type and volume of fire

Kill zones of fire

Orientation (aiming) of the weapons

Operational mode of friendly weapons (destroyed, silent, reloading, firing)

Condition of enemy equipment (operable, destroyed)

Status of ammunition supply

In addition, the simulation should reflect the effects of weapons activity on visibility on the battlefield and inside buildings.

Unless DIs are dealing with their own weapons, the visual cues will be located some distance away. Even when operating a weapon, unless the DI is in a close combat situation as discussed above, rapid scanning movements will not be required to perform this assessment. Level 1 technology is thus considered adequate to meet this simulation requirement.

4.1.11 Assess Other Battlefield Details

The DI must see and deal with a host of battlefield details beyond those discussed above. A representative sample of such details includes:

Various supplies, including ones that are air-dropped

Mines, which must be detected and then assessed in terms of their operational status

Barbed wire and pickets

Trip wires for mines, which must be assessed in terms of their tautness and integrity

Small obstacles

Communications lines

Conditions in the trench

Status of obstacles that have been breached (e.g., area of breach)

Miscellaneous equipment (ropes, grappling hooks, litters, wire & bolt cutters, etc.)

Evidence of enemy activity such as tracks, litter, obstacles, booby traps.

These objects must be seen, and some must be physically handled as well in the operational setting. For the most part, these objects will be located close at hand but are not likely to require large and rapid scanning motions. In addition to the potential computational burden imposed by the need to display these objects, there is some question as to the level of display technology required for the soldier/simulator interface. We have indicated that Level 1 technology is considered adequate for all training conditions, but that some additional training benefit might be realized by employing the HMD (Level 2 or 3) for MST training where it is more important for the trainee to have a higher scanning bandwidth.

4.1.12 Use Navigation Aids

In addition to performing navigation by viewing the visual scene, the DI will use a compass and other aids such as maps and navigation sheets. Level 1 technology is adequate for displaying a computer-generated image of a compass. (Part of one screen could be devoted to this purpose). Navigation aids other than the compass could be provided by using the real objects, or by temporarily devoting a screen or part of a screen to suitable electronic presentations.

The method of providing a particular type of navigation information with Level 1 technology will depend, in part, on the amount of detailed information required. For example, it should be feasible to provide map-like information of modest information density and small area on the screen, similar to what is now being developed for automobile electronic navigation aids. On the other hand, map information of high density and/or large area would likely require a paper map. (Note that all such visual information must be provided electronically when the soldiers uses a Level 2 or 3 HMD.)

4.1.13 Computational Requirements

The preceding discussion has focussed on the technology related to the soldier/simulator interface, i.e. the display device. Of course, in order to realize the training potential of the ICS, it is critical that the system be able to provide the

necessary information that is to be displayed. This requires that (1) the simulator contains or generates the required information base, and (2) that there is sufficient computing power to generate the scenes in real time with sufficient detail and acceptably low delay. The CCTT specifications, in effect, requires this computational capability, and our best guess is that, given the trends in computer technology, this capability will realized during the period of CCTT development.

4.2 Visual Sensing

Visual sensing by the simulator, as we have defined it, is required only for HMD visual presentation. The simulated combat situations requiring Level 2 display technology (urban fighting and other close-in situations) therefore require visual sensing technology of Level 2 or higher (sensing of head, finger, and hand position and movement).

For the initial implementation of the CCTT, troops will be commanded either by radio/telephone (communication between trainees) or by joystick (to command simulated troops). Because hand and arm signals are often used in field operations, the capability for allowing the simulator to sense this activity has been considered desirable for MST, where it is most important for the trainee to get a feel for the actual operation. Level 2 and higher visual sensing technology is required to implement this capability.

4.3 Auditory Display

The CCTT is required to provide battlefield sounds using loudspeakers only (headsets are to be used only for communication). Given this requirement, plus the current use of speaker systems in the SIMNET battlefield simulation, use of speakers alone for providing battlefield sounds is defined as Level 1 technology. Combined use of headphones and speakers for battlefield sounds is considered Level 2.

As noted in Chapter 3, use of Level 2 technology is recommended when it is important for the trainee to localize sounds that are likely to be produced nearby from any direction, or when the acoustic environment must be replicated with especially high accuracy. Level 1 technology is recommended otherwise.

4.3.1 Communication Among Trainees

Communication at a distance will be conducted via radio/telephone, but the simulation should allow direct verbal communication among trainees within (simulated) hearing distance. Level 2 is recommended for MST (and perhaps the advanced stages

of CPT) where precise localization of the individual speaking is critical; Level 1 is recommended otherwise.

4.3.2 Conduct Operations in Urban or Close-in Environments

A definition of "operations in an urban environment", and the associated demands on visual information requirements and VET requirements is given in Section 4.1.4. Because the trainee will want to localize sounds from nearby sources located in any direction, Level 2 or 3 technology is recommended for the more critical MST exercises. Level 1 is probably adequate for the more cognitive CPT and MPR operations, but some training enhancement is likely to be provided by Level 2 or 3.

The same considerations apply to non-urban close-in environments.

4.3.3 Determine Status of Enemy and Friendly Vehicles

Although visual cues are expected to play the major role in determining the status of enemy and friendly vehicles, relevant information can also be gained from engine, transmission, and other sounds made by the vehicles. Because this sort of judgment does not depend on precise auditory localization of the sounds, Level 1 technology is considered adequate.

4.3.4 Assess Status and Effects of Weapons

For weapons operating beyond line of sight - especially hostile fire - auditory cues may be the only source of certain types of information (e.g., the rate of firing of various large weapons, the use of other explosives, etc.). Because auditory cues of this sort are associated with action at a distance, Level 1 technology is adequate for performing localization and identification.

4.3.5 Be Subjected to Disruptive Effects of Battlefield Noise

We have not considered it important to represent the general stress effects of battlefield sounds for the types of training considered in this report. On the other hand, should battlefield sounds impede the activities of the trained soldier to the point that mission effectiveness is impaired, such effects must be accounted for to allow realistic mission planning, rehearsal, and training.

Aside from inducing anxiety from constant exposure to loud noises, the acoustic environment is most likely to degrade mission success because of the interference by one set of sounds with the acquisition of information from another set of sounds. Examples include interference with voice communication, either direct or via radio/telephone, masking of sounds of enemy

movement, and interference with the ability to detect, identify, and localize specific enemy weaponry.

Level 1 technology is considered adequate for replicating these type of disruptive effects.

4.4 Auditory Sensing

Recall that "auditory sensing" refers to the capability of the ICS to accept spoken language from the trainee and extract information from the spoken words so as to influence the behavior of the simulated DI (i.e., allow the trainee to verbally command simulated troops and/or cause the simulator to take other actions). The capability to allow one trainee to communicate with another is assumed to be available and is not considered an aspect of VETT.

Because the baseline simulator configuration will not have this capability, Level 1 technology is defined as the absence of speech recognition.

The capability to command troops verbally is considered desirable (but not necessary) for MST (and perhaps advanced stages of CPT) where physical realism is more of an issue, and not required for CPT or MPR. Because of the limited number of trainees that would use a given simulator host at any one time, and because a relatively limited vocabulary is expected to be sufficient, Level 2 technology (limited speech recognition) is recommended. If available and cost effective, Level 3 technology (advanced speech recognition) is preferred.

4.5 Haptic Display

For the most part, we consider desirable, but not necessary, the presentation through haptic mechanisms of information concerning the state of the world or of the physical condition of the DI. The levels of haptic display technology recommended for meeting specific information needs are given below.

4.5.1 Receive Feedback Concerning Physical Condition of the Troops

For CPT, which tends to train procedural and tactical skills for relatively short-term tasks, the physical condition of the simulated DI is not likely to be an issue (provided the soldier is not a battle casualty in the sense of being killed or wounded). On the other hand, accurate planning and realistic training for a specific, relatively long-term mission should account for the potential degradation in fighting capability of the DI due to physical stress. To the extent that it is desirable to indicate to trainees the degradation in mobility of the DI(s) they command, Level 2 technology (e.g., control devices

with programmable force/displacement characteristics) is appropriate for providing physical feedback related to the performance capabilities of the DI.

Other modalities for indicating the effects of performance degradation are discussed in Section 4.9.

4.5.2 Realistic Feedback From Control and Manipulation of Weapons, Etc.

Although the type of training considered in this report is largely procedural and is not intended to develop primary physical skills, realistic simulation of a military mission involving the DI will generally include a number of activities that require (simulated) physical interaction with battlefield objects. Such activities include aiming and firing hand-held and crew-served weapons, operating specialized equipment, handling mines, installing or cutting barbed wire, breaching objects, etc.

Level 3 technology (programmable general-purpose devices) is required to provide this type of physical feedback. As discussed in Chapter 5, significant advances in the state of the art are needed to make this technology practical for ICS applications.

4.5.3 Assess Battlefield Conditions

Although assessment of battlefield conditions is largely through the visual sense, there are some conditions where the sense of touch provides a useful augmentation (and may even be required to effectively assess the situation). Some of the activities that generally require physical interaction for proper assessment are: (1) analyzing surface conditions with regard to suitability as a landing site or as a path for transporting heavy equipment; (2) estimating the tautness of a trip wire; and (3) evaluating the integrity of barbed wire installations and other impediments to movement.

The ICS will presumably have the capability to present information of this sort by visual means using shape, color, shading, and/or other types of visual coding. This artifice is sufficient for planning purposes, where it is sufficient to reflect the outcome of the observer's assessment without recreating the mechanism by which the assessment is made. (For example, if a particular surface area is unsuitable for heavy operations, it is assumed that the observer would correctly reach this conclusion, and the area is displayed with a particular visual texture that indicates unsuitability.)

For MST, and perhaps for CPT as well, the DI's training may be enhanced (and the outcome of the operation more accurately predicted) if the assessment is done in the more operationally faithful manner through the sense of touch. Because of the

variety of touch-related stimuli that would be needed to provide proper cuing for the range of activities in this category, Level 3 technology is required if haptic information of this sort is to be provided.

4.5.4 Encumbrance of Protective and Other Gear

Certain tasks or battlefield environments will require the DI to wear protective clothing ranging from single items such as masks or gloves to fully-protective suits for NBC protection. The use of such gear can degrade performance in a variety of ways: gloves can hinder the manipulation of switches and buttons as well as the handling of small objects; masks will limit the FOV and may fog; full protective suits can impede movement and cause heat stress. Carrying particularly heavy packs, large pieces of equipment, or other heavy objects may also impede movement.

The technology (or combination of technologies) appropriate for reflecting this cause of performance degradation depends on the particular effect that is to be induced or simulated. One approach to simulating at least some of the effects of protective clothing is to have the trainee actually wear the item(s). A properly instrumented face mask could be made to fog, or a suit could be heated, whenever the simulated activity of the trainee was such that, operationally, the mask would likely fog or the DI would likely become overheated.

If the trainee is sitting or standing at a workstation - the mode of operation envisioned for the foreseeable future - simply wearing the gear (or carrying the object) would not limit mobility with respect to the simulated battlefield activity. As suggested in Chapter 3, Level 2 haptic display technology (programmable force/displacement control device) coupled with software limitations on the rate of travel of the DI icon is recommended for representing this type of performance degradation.

4.6 Haptic Sensing

The trainee must provide certain command information to the ICS though some form of manual control input. In keeping with the terminology used so far, the mode by which the simulator receives this information is referred to as "haptic sensing". Control inputs relevant to battlefield operations (as opposed to the operation of the simulator itself) include simulating movement of the trainee (and/or other simulated DI) over the battlefield, and performing other manual operations such as aiming and firing weapons, using other equipment, handling objects, etc.

Because the main purpose of the ICS is to train procedural and tactical skills, rather than physical skills, Level 1 technology (standard input devices such as joysticks, switch panels, and touch screens) is considered adequate to meet training needs. However, if Level 3 haptic display technology (programmable general purpose devices) is used in order to provide the desired haptic feedback as discussed above, then, since the same device provides both input to and output from the ICS, Level 3 technology will also be used for providing the control input.

4.7 Whole-Body Movement

Issues relating to the simulation and perception of whole-body movement are discussed in Section 3.3.7.

4.8 DI Models: Biomechanical Articulation

Level 2 modeling technology (indicating arm and limb positions of the DI) is recommended for certain battle environments as described below; implementation of Level 2 technology, however, is feasible only if a DI icon corresponds to a single (real or simulated) soldier. In the present SIMNET implementation, a single icon represents an entire squad. Initial implementation of the CCTT is expected to use a single icon to represent a fire team (i.e., about half a squad). Because the position of head and limbs is highly individualistic, Level 1 technology seems most appropriate for representing the aggregate behavior of a group of individuals.

Level 2 modeling of the DI is most appropriate when (1) operations are conducted in an urban or other close-in environment that requires soldiers to operate near one another; (2) Level 2 or 3 technology (HMD) is used for the visual display (generally a requirement imposed by condition 1); and (3) each DI icon represents a (live) trainee, rather than a simulated DI.

Close physical interaction with a fellow soldier requires a more detailed indication of the status and activity of the other soldier than is available with Level 1 (fixed icon) modeling. For example, a trainee who wants to point out or show an object to another soldier must first ascertain where the other soldier is looking in order to judge the effectiveness of the communication. The trainee may also need to know other aspects of other soldiers' activities, such as whether or not their weapons are in position to be fired. This level of DI modeling only makes most sense if the DI icon represents another trainee having independent modes of operation; otherwise, the icon represents a computer model of a DI most likely under the control of the trainee, in which case (1) the trainee knows what the DI is supposed to be doing, and (2) the need to command head and

limb motions in addition to translation and orientation of the body would impose excessive workload on the trainee.

The applicability of Level 3 technology (fully animated DI models) is the same as that for Level 2. Once appropriate algorithms have been developed for animating the DI model, Level 3 will be preferred over Level 2 in cases where there is sufficient computing power to support full DI animation.

4.9 DI Models: Influence of Physical Condition

As discussed in Section 3.2.10, three levels of DI modeling are considered for representing degraded physical condition of the simulated DI due to physical stress such as fatigue, thirst, etc. (as opposed to performance degradations discussed above related to physical encumbrances). These levels are: (1) indication only of killed, wounded, or operational status; (2) for an operational DI, reduced attainable movement speed; and (3), in addition to reduced mobility, a modification of the physical appearance of the DI.

Level 1 modeling is recommended for CPT, where the physical condition of the soldier is not likely to be an issue. Level 2 or 3 is recommended for MPR and MST, where accounting for the effects of fatigue and other sources of performance decrement are important both for accurately predicting the outcome of a military operation a well as to provide "honest" training for the squad leader whereby the effects of physical limitations are understood and either avoided or compensated for.

4.10 Physical Condition of the Trainee

It is as important for trainees to be made aware of the effects of stress on themselves (i.e., what the effects would be if they were in the actual combat situation) as well as the effects of stress on simulated troops under their command. The icon representing the trainee to others would be treated as described above. The issue here is to induce a physical feeling, through some form of artificial stimulation, that is representative of the stress computed by the simulator.

No special technology is recommended for CPT, whereas Level 2 technology (movement speed degraded as appropriate) or Level 3 technology (Level 2 plus artifically induced stress) is recommended for MPR and MST.

As suggested in Section 3.2.10, Levels 2 and 3 are needed mainly when trainees represent only themselves, rather than an entire unit; otherwise, the treatment discussed above for the simulated DI applies.

Intentionally left blank

5. Feasibility of Using VETT for ICS

Level 1 technology, as we have defined it in the preceding chapters, is generally available at present for application to ICS. Additional hardware and software developments are needed, however, to make the more advanced VE technology feasible for meeting combined arms training needs. These requirements are discussed below, after we first review the rationale for applying VETT specifically to ICS training requirements.

5.1. ICS as a First Priority Application of VETT

The Statement of Work suggested the need to assess the differences in feasibility and, where possible, cost of using VETT for ICS when contrasted to using it for other Army applications such as maintenance training, flight training, or weapons system operator training. This is a difficult task because the range of sophistication of potential application in any of these domains is so broad as to deny generalization. Alternatively we have chosen to identify some of the arguments for making ICS a first priority application.

The ICS application is in many ways unique. It is unique in requirements, particularly the requirement to reflect multiple individuals moving on a battlefield or urban environment interacting with each other and with other combined arms units. It requires attention to whole body movement to an extent not emphasized in other applications, particularly flight training applications. The need, ultimately to reflect whole body interaction with other physical objects such as climbing stairs, climbing in windows, searching buildings, digging trenches, etc., will push the haptic simulation capabilities to their limits, if they are realizable at all. These issues are sufficiently difficult that they will force the development of alternatives to full "artificial reality", and that may be laudatory.

The ICS is a development that will provide an immediate stimulus to drive the cost of VET down. Each individual soldier that is represented as an individual will require a helmet-mounted display, head position sensors, acoustic displays, a whole body movement control system, and haptic sensors and display to the extent that they are available. This application is not like training a pilot to fly a multi-million dollar aircraft for which a large investment in training simulators is justified. It is hard to imagine equipping very many individual soldiers with training devices costing more than $50,000 per soldier to support small unit or even combined arms training on the battlefield. There is interest in low cost systems for the entertainment market, but the ICS application will provide a healthy stimulus for cost competition for a more sophisticated system than can be justified for the entertainment market.

Combined arms training in SIMNET already represents a step in the direction of virtual environments. The CCTT procurement specifies that there will be workstations representing ICS, at least at the level of a platoon or squad. This becomes the natural starting point on which to build more sophisticated VEs. Thus, CCTT provides a path for graceful upgrade from a minimal system to sophisticated representation of collections of ICSs. Our three levels of development are suggestive of such an upgrade path.

It seems likely that there are classes of team coordination training that can be done by VET that are difficult to accomplish in any other way, short of expensive full-scale field exercises with actual warfighting equipment. SIMNET has found combined arms training to provide unique readiness skills reflecting the ability to coordinate units and understand each other's perspectives. It is likely that these opportunities for individual soldiers to train coordination skills with other units and with each other can be accomplished more satisfactorily using ICS.

Finally, simulation of land combat is of interest primarily to the Army. It is unlikely that the unique aspects of the technology required to accomplish ICS will be developed by the other services. It reflects a requirement that will not be satisfied or even pursued by the other services. If the Army wants the capability and, judging from CCTT it does, then the Army will have to invest in its development.

5.2 Required Technology Development

Major hardware and software developments required to make ICS feasible for training dismounted infantry are discussed. As implied by the material in the preceding chapters, we believe that technology employed in the current SIMNET DI simulation, along with other readily available conventional technology, can meet many of the procedural and tactical training requirements of CPT and MPR and, to some extent, MST as well. More advanced technology, much of it associated with providing head-mounted virtual environments, is needed to handle battle environments that require the DI to respond rapidly to events occurring around him. Advanced technology is desired for MST to allow the DI to perform tasks in a more realistic manner than can be provided by conventional technology.

The discussion of technology requirements is organized by sensory modality and parallels the organization of Table 2. Needs for hardware and software developments, if relevant, are summarized and we offer estimates as to when these developments will be realized.

Cost of equipment has not been considered in the preceding chapters of this report. We must, however, consider cost factors when assessing the feasibility of applying VE technology to military training systems. In the discussion that follows, estimates of the "availability" imply availability at affordable cost. While a consensus on the definition of "affordability" remains to be achieved, we have suggested $50,000 as an upper limit on the cost of an individual DI workstation in the ICS. To meet this cost goal, major components should not cost more than $5,000 to $10,000 each. We consider the criterion cost that distinguishes "affordable" from "high-cost" to lie within this range.

This chapter deals with the state of technology in relatively general terms. Readers interested in the costs, availability, and technical details are directed to Durlach et al. (1992).

5.2.1 Visual Display

Multi-screen displays are currently in use for the SIMNET DI simulation. No significant hardware or software developments are required to employ this technology.

Level 2 visual display technology is defined as a helmet-mounted display (HMD) having capabilities adequate for most training purposes at an affordable cost. Level 3 technology approaches the limits of human visual perceptual capabilities sufficiently so that device limitations have no deleterious impact on training.

A Level 2 HMD will typically have uniform resolution over the device's entire FOV. A Level 3 device may have a uniform high resolution over the entire FOV, or a high-resolution central area surrounded by a lower resolution peripheral area. Unless the high-resolution area covers a field of view as great as that typically scanned by the human without making head movements, eye movement tracking will be required so that the high resolution segment is roughly centered on the operator's point of regard.

In the absence of data to the contrary, we assume that, to provide a visual scene suitable for training purposes, the HMD should provide (1) color, (2) a FOV comparable to that of the human eye, and (3) a resolution comparable to that of the human eye, at least in the central viewing area. The visual presentation must be reasonably free of distortions such as delay, poor focus, geometric distortion, low contrast ratio, etc., and, of course, the device must be safe to use.

Currently available devices do not meet these rather stringent requirements, and even Level 2 devices tend to be relatively expensive. Development is proceeding at a rapid pace,

and we expect Level 2 (and possibly Level 3) devices to be available for training systems within the 3-5 year frame.

5.2.2 Visual Sensing

Considerable activity is underway to develop electromagnetic and optical devices for tracking body position (especially head, hand, and finger position). Some adequate but relatively expensive systems are available at present. We anticipate that good systems at affordable prices will be available within a few years.

Eye position trackers have been in existence for some time. Devices suitable for VE applications tend to be relatively expensive, however. In addition, the need for careful and repeated calibration, and the difficulty of co-locating an eye-movement tracker with a head position sensing device leaves this technology somewhat problematical. Continuing developments of high-resolution, wide FOV HMDs may well eliminate the need for eye trackers for training in the ICS.

5.2.3 Auditory Display

Level 1 auditory display technology is limited not so much by deficiencies in speaker technology but more by the problems associated with the use of speaker arrays. For example, considerable spatial volume is required to surround each trainee with a speaker array in order to maintain reasonable distance between the trainee and the speakers. Adequate acoustic shielding of co-located DI stations could prove to be a significant expense.

Because of these and other limitations (See Durlach et al., 1992), advances in auditory displays are expected to focus on Level 2 technology, which we envision as using a combination of headphones and speakers to present battlefield sounds. The headphones would serve as the primary display device, with the speakers used for providing very low frequency sounds that tend to induce whole body vibrations (e.g., earth-shaking explosions).

Because very low frequency sounds are not used for localization, a single speaker is sufficient for a DI station, and shielding may not be needed between co-located DI stations for which the operators were playing the roles of soldiers that would be in close proximity in the actual battle.

A device such as the Convolvotron that accounts for the effects of the human's physical structures on the perception of sound - based on measurements of head-related transfer functions (HRTFs) - appears most promising for VE application. Certain technical limitations need to be overcome to make this type of device feasible for ICS training, including the handling of

reverberant environments and better techniques for interpolating among the stored HRTFs. We expect this to occur within 2-3 years.

5.2.4 Auditory Sensing

Factors that determine the availability and cost of speech recognition systems include (1) the extent to which the system is speaker dependent; (2) the type of speech recognized, ranging from discrete words or phrases (substantial pauses between words) to continuous speech; (3) size of the allowable vocabulary, (4) grammatical constraints, and (5) allowable background noise.

We assume that for the kind of command and control environment posed by the ICS, the following characteristics of a speech recognition system would be adequate for use in the ICS:

Able to recognize discrete words and short phrases

Relatively small vocabulary

Very tight grammatical constraints

Able to operate in a noisy environment

Systems whose specs meet these requirements - which we would classify as Level 2 technology - are currently available at costs of under $2000 and can operate in either a speaker-dependent or speaker-independent mode.

Substantial advances in speech recognition and speech understanding are expected over the next few years, along with the amount of computing power available at a given cost. Whether these advances will be sufficient to provide Level 3 speech recognition technology at an affordable cost remains problematic, however.

5.2.5 Haptic Devices

Control devices with programmable force/displacement characteristics are currently available and are being refined over time. Level 2 haptic display and sensing capabilities thus appear feasible over the near term.

Level 3 technology - programmable general-purpose devices - fall largely into two categories: exoskeletons (grounded either to the human or to the environment), and tactile displays. Development of tactile displays is continuing. We expect useable tactile displays having relatively coarse resolution to be available within 2-3 years. Substantially longer time will be required to develop tactile displays that match the resolution of the human's haptic sensory system.

Development of exoskeletal devices (beyond the level of the data glove) is an active area of research - one that is not sufficiently mature for us to speculate on the time required to provide devices feasible for ICS application.

5.2.6 Whole Body Movement

We have defined Level 2 technology as the simulation of large-volume movement through mechanisms that allow movement in place. We are unaware of any substantive research in this area and we therefore cannot reasonably speculate as to the time when adequate capability will be available at affordable cost (or whether such devices will be developed at all). Because technology suitable for the ICS is not covered in detail in Durlach et al., (1992), we review some of the technical requirements and critique two approaches relevant to whole-body movement.

A whole-body motion cuing device potentially serves two purposes: it provides cues to the simulator as to the intended movement of the trainee, and it provides the sensation of body movement to the trainee. The latter capability is especially relevant when the trainee is using a HMD, where good correspondence between whole-body motion cuing and visual cuing of self-motion may help avoid or alleviate systems associated with simulator sickness. In order to limit the real estate requirements of the ICS, we impose the constraint that actual physical movements occur within limited volume.

Ideally, a whole-body motion cuing device will meet the following requirements:

Provide inputs to the simulator regarding movement intention

Allow trainees to move their limbs in a natural fashion

Provide appropriate acceleration cues to the DI

Accommodate motion in three translational dimensions

Allow the trainee to assume different stances and to change orientation

Provide feedback appropriate to the terrain (e.g., the feeling of climbing a hill)

Confine actual physical movements to within a small volume

Given that we do not allow the trainee to cover a large amount of territory while participating in the simulation, these requirements are to some extent self-contradictory, and different approaches tend to accommodate different requirements.

One potential approach is to have the trainee operate on a two-dimensional treadmill while using a HMD display and some type of exoskeleton to provide a virtual control and display environment. The treadmill would allow limb movement appropriate to movement on the surface of the earth; it could be tilted to represent non-level terrain. It would not be adequate for simulating stair climbing, however. In order to provide translational acceleration cues, the treadmill surface available to the trainee would have to be large enough to allow some actual translation, and some sort of "washout" (position restoring algorithm) would be needed to prevent the trainee from leaving the treadmill.

A second approach is to provide an earth-grounded exoskeleton that operates in three dimensions so that stair-climbing as well as movement along the terrain can be simulated. Again, some form of washout system would be required to constrain actual physical translation.

Designing the control systems for either of these devices is expected to pose a significant technical challenge.

5.2.7 DI Models: Biomechanical Articulation

The DI model currently implemented in SIMNET shows a perspective view of the icon appropriate to stance (standing, kneeling, or prone). Articulation of head and limbs is not currently implemented. No technical breakthroughs are needed to achieve the technology to display an articulated DI icon, other than the availability of sufficient computing capability to accommodate the higher degree of freedom DI models. On the other hand, considerable effort may be required to develop algorithms to drive the DI models to achieve motions that look natural, particularly when the DI icon interacts with objects in the environment.

5.2.8 DI Models: Influence of Physical Condition

Representation of the (simulated) DI's physical condition as suggested in this report does not require any specific technology development other than the further development of DI modeling. The following capabilities are suggested for implementation:

Modify the displayed image to allow the trainee to differentiate between killed, wounded, and operational. Modify squad firepower appropriately.

Develop and implement algorithms for determining the physical condition of the simulated, functional, DI (e.g., degree of fatigue), and adjust the allowable movement speed of the DI accordingly. If the speed is controlled by the trainee through a joystick, adjust the force/displacement

characteristics to reflect the fact that more effort is require to achieve a given movement speed.

Modify the appearance of the functional DI icon to reflect physical condition.

5.2.9 Physical Condition of the Trainee

To the best of our knowledge the notion of indicating to the trainees what their physical condition would be in the operational situation has not heretofore been seriously considered. Nevertheless, we feel that implementation of this capability would allow for more realistic mission planning and training.

Level 2 technology is the same as for the computerized DI model; namely, modify the allowable movement speed (and adjust force/displacement characteristics of the joystick if relevant) to reflect a degradation of physical condition. Level 3 technology, which involves subjecting the trainee to real (but situationally artificial) stress, is an area for new research; it is premature at this time to speculate as to when this capability would be available for use in the ICS.

6. Research Tasks and Facility Requirements

Much of the research recommended in Durlach et al. (1992) for Navy training needs applies to the application of virtual environment technology to Army training requirements. The reader is referred to Chapter IV of that report for more detailed development of some of the ideas presented below and for additional research topics, many of which will apply to Army needs.

Because the focus of this report is on the application of VE technology to the training of dismounted infantry, the research areas outlined below are aimed at exploring the utility of VET specifically for training purposes. The tests of the technology will therefore be in terms of "performance fidelity" and "training fidelity" which, as we discuss below, are related but not identical concepts. In general, different training purposes tend to impose different fidelity requirements.

Full "performance fidelity" is achieved when the soldier performs a task in the simulator with the same result as when the task is performed in the real world, according to any objective measurements of human and system performance. Because we can only measure overt behavior, we generally cannot verify that the human is using the same informational quantities, even when the simulated and operational stimulus environments appear to be similar. Performance fidelity is required primarily for mission planning where accurate estimations of time requirements and outcomes are important. For planning purposes, it is as important that simulator performance be no better than operational performance, as it is that simulator performance be no less effective.

Full "training fidelity" is achieved when the training provided by the simulator enables the soldier to perform a real-world task as well or better than if all training had been in the operational situation. Training fidelity is needed for all training purposes other than planning; specifically, combat proficiency training, mission rehearsal, and mission-specific training.

Training fidelity need not always imply performance fidelity. For example, if the information environment is in one or more ways deficient compared to the operational environment, performance will most likely be degraded compared to the operational task. Nevertheless, the greater information-processing workload imposed by this situation may make the soldier a more efficient user of the information, with the result that performance achieved upon transfer to the operational situation will equal or exceed what would have been elicited through real-world training.

Conversely, performance fidelity need not always imply training fidelity. If the trainee uses different perceptual cues in the simulator (or uses the same cues but in a different manner), performance may approximate that produced by a trained soldier performing the operational task, but the information-processing strategies may not transfer well from the simulator to the real world.

The remainder of this chapter is organized into two sections. The first recommends areas for further research, and the second outlines the structure of a research facility with which to conduct the proposed research.

6.1 Areas for Research

Areas for further research are organized into two categories: (1) studies that focus on training techniques and procedures relevant to training in a virtual environment, and (2) studies related to exploring requirements for specific sensory systems, devices, and procedures that are associated with VETT. Many of the issues discussed here apply to combined arms training in general and are not specific to the DI.

6.1.1 Training Issues

Five topics are addressed in this section: (1) methods for measuring transfer of training, (2) exploring the training effectiveness of Level 1 technology as defined in the preceding chapters of this report, (3) the benefits of artificial or enhanced cues, (4) simulator sickness, and (5) presence.

6.1.1.1 Methods for Measuring Transfer of Training

Because the intended application of virtual environment technology considered in this report is for training the dismounted infantryman in the Individual Combat Simulator, the VE technology is most appropriately evaluated with regard to training effectiveness. Accordingly, we recommend that studies be conducted to determine the appropriate methods for comparing performance in the simulator and in the field (test for performance fidelity) and for assessing the transfer of skills from the simulator to the operational situation (training fidelity). Development of testing procedures of this sort is vital to the success of the research recommended below in Section 6.1.2.

Definitive transfer-of-training studies are difficult to design and conduct. Because a given subject can only be trained one way for a particular task, and processed only once through a given training program, large numbers of subjects are needed to provide statistically useable data. Furthermore, it is not sufficient to simply compare performance across the different

subject populations; one must also consider training time and, ideally, the way in which task proficiency is achieved. Issues associated with the measurement and prediction of transfer of training are discussed in Chapter III-B-2 of Durlach et al. (1992).

To study transfer of training provided to the DI by the ICS, tasks, performance measures, and performance criteria must be carefully defined. The ARTEPs that are currently used to specify DI training will provide a solid basis for defining these parameters.

6.1.1.2 Determine Training Effectiveness of Level 1 Technology

We have suggested in earlier chapters that most of the training requirements of the ICS can be met with technology currently existing and projected for the near term (defined as "Level 1" technology). To the extent this is so, significant savings in time and equipment costs can be realized, and useful training can commence while waiting for the more sophisticated technology to become available and cost effective. Furthermore, as we note below, incidence of simulator sickness is expected to be less with Level 1 technology than with other technologies that enhance the feeling of "presence."

We therefore recommend that a research program first be undertaken to determine the extent to which useful training of the DI can be accomplished with Level 1 technology. The first study of this program compares CPT in the ICS (using Level 1 technology) with current methods. Different groups of soldiers are initially trained by these two methods and then subjected to the same proficiency test. On the basis of these results, one of these two training methods is tentatively selected as being the more cost effective (i.e., the method that would most likely be adopted if one had to choose based on this study). In a subsequent series of training transfer studies, the method selected from the first study is compared with ICS training using Level 2 technology.

The results of this study program will provide a basis for comparing the cost effectiveness of "standard", Level 1 ICS, and Level 2 ICS training schemes and thus provide a basis for determining future Army training needs.

6.1.1.3 Benefits of Enhanced and Artificial Cuing

Simulation using VET has the potential to provide enhanced or artificial perceptual cues to increase training effectiveness. These cues either do not exist in the real world, or are modifications of cues that do exist in the real world. This issue applies to training in general and, as we show below, to DI training.

We consider four categories of enhanced cuing: (1) indication of pedagogical intent, (2) compensation for physical limitations of the cuing devices, (3) intentional improvement of the cuing environment to speed up the learning process, and (4) intentional degradation of the cuing environment to induce the trainees to develop a higher level of skill than they would otherwise. For reasons suggested below, considerable research is needed for the latter three categories of cue enhancement.

The first type of cuing is equivalent to an instructor pointing to the object under discussion. For example, if the DI's mission is to secure a building, the displayed image of this building might initially be highlighted by some means (flashing, bright color, indicated by an arrow, etc.) to indicate to the DI which building is to be secured. Once the maneuver begins, the building is then displayed with its "normal" (at least, not intentionally distorted) appearance. The benefits of this type of enhancement seem obvious. A demonstration study should suffice; we do not foresee the need for substantial research into this type of cuing.

To some extent, the second type of cuing - compensation for simulator limitations - is expected to be employed substantially. Such compensation is required when DIs are required to perform a task in the simulator in an entirely different manner than they would in the real world. For example, the use of a weapon might be initiated by the trainee by depressing a switch. The simulator then calls upon a model to determine the time that would be consumed in this operation, plus the likely effect of the weapon on its intended target. Ideally, the statistics for time required and damage inflicted will be such as to maintain a limited type of performance fidelity with respect to the real world (i.e., system performance is replicated even though operator behavior is different).

Compensation may also be used to maintain cuing effectiveness in the presence of display limitations. For example, suppose that a particular landmark is generally detectable at a distance of 5 km in the real world, but, because of display resolution limitations, is not detectable until a distance of 2 km when displayed in proper perspective. One might try to maintain performance equivalence by distorting the cue (say, display with a larger-than-real visual size) so that the landmark is detected at 5 km in the simulator. There are a number of problems with this type of cue distortion. Techniques of this sort have been tried, but to our knowledge their training effectiveness has not been established, and this type of cue enhancement remains a potential research issue.

Even if performance equivalence is achieved, there is no guarantee of adequate training fidelity. This is of particular concern where artificial cuing is employed to compensate for

display limitations, where the trainee may learn to rely on the enhanced cue to the detriment of learning to use the cues that would be available in the real world. In the example cited above, if trainees are provided with the sudden appearance of an image that is immediately recognizable, they may be delayed in learning to recognize the landmark in the real world as it proceeds gradually from an apparent speck on the horizon to a recognizable image.

The third category of cuing differs from the second in that the simulated display environment is in some way made better than the operational display environment in the hopes of speeding up training. For example, let us assume that the differentiation between enemy and friendly vehicles seen at a distance is made on the basis of shape. One potential way to speed up training in the simulator is to exaggerate the shape differences early in training and gradually present more realistic shapes as training proceeds. It seems clear that vehicle detection performance will be better early in training with the exaggerated shapes than with the realistic shapes, but it is not clear that the DI will train faster to identify vehicles that are portrayed realistically. It is possible that the presentation of exaggerated shapes retards the ability to distinguish among realistically-portrayed vehicles, in which case artificial cuing will have slowed rather than enhanced the learning process.

The fourth category is similar to the third in that the cuing environment is distorted not to compensate for simulator inadequacies but to improve training - in this case, to enhance skill development. The difference is that the perceptual environment is intentionally degraded rather than enhanced in accordance with the theory that if humans can learn to perform in a degraded task environment, they should be able to perform exceptionally well in the operational environment.

Although the notion of enhancing or otherwise modifying the cuing environment to improve training is an appealing one, it involves the considerable risk that distortion of the cuing environment will induce a false reliance on artificial cues or otherwise induce an inappropriate response strategy that tends to retard the trainee's ability to function in the operational environment.

We therefore recommend that a series of studies be undertaken to determine the conditions under which the various types of cue enhancement improve transfer of training.

6.1.1.4 Simulator Sickness

One of the paradoxes facing the application of VETT is that the better we make the visual system, the more likely the case that trainees will be unable to use the system because of a

phenomenon known as "simulator sickness." Simulator sickness may be defined as a feeling of discomfort that arises from performing tasks in the simulator, where such discomfort is not elicited when the same tasks are performed operationally. This discomfort may include nausea and disorientation that occur while the simulated tasks are being performed, plus adverse symptoms that persist (or become initially apparent) after the person has left the simulator.

Simulator sickness is believed to be caused, in part, by sensory cue conflict in which the motions perceived by one sensory mechanism are not confirmed by the motions (or lack thereof) perceived by another sensory system. This effect is generally elicited in a fixed-base (stationary) simulator in which the visual perception of motion is not confirmed by whole-body motion cues, or in a moving-base simulator having poor motion cuing fidelity. The more compelling the visual scene, the more likely, and more intense, the induced "sickness." The HMD is expected to provide a more compelling visual scene (a sense of "presence") and thus be more likely to induce simulator sickness than, say, an array of display screens having a lower FOV.

We differentiate between simulator sickness and motion sickness in the sense that motion sickness consists of discomfort that accompanies actual self-motion. If motion sickness would occur when performing a task in the real world, than a faithful simulation should elicit this effect. Simulator sickness, however, is always an undesirable effect, because it represents a negative effect that is (by definition) not present in the real world.

Progress has been made in defining the relationship between simulator sickness and the details of the task being simulated, but methods for overcoming simulator sickness need further development. In particular, research is required to determine the best way to deal with simulator sickness in the context of ICS training. Possible approaches include (1) select out trainees who, by some testing procedure, are considered likely to be susceptible to simulator sickness; (2) design the simulation to avoid sickness; and (3) allow trainees time to adapt and overcome the tendency to sickness.

Selecting out trainees may be appropriate for studies designed to explore training issues, but is not appropriate for operational use of the ICS, which is to train soldiers who are otherwise qualified for the tasks in which they are to be trained.

Use of Level 1 technology (which does not call for the HMD) is one way to minimize the incidence of simulator sickness, as the sense of presence is expected to be less compelling than with the use of more advanced display technologies. If, as we expect,

there will be some tasks requiring the HMD, other methods for reducing sickness should be explored. This might involve reducing fidelity (e.g., blanking or temporarily reducing the FOV) to lessen the impact of sensory cue mismatch, or providing additional cuing such as limited whole-body motion.

Adaptation is another approach, whereby the trainee is given repeated exposure to the simulation until in effect, an "immunity" to simulator sickness is built up. The problem with this approach, other than the additional training time required, is that trainees may adapt by taking certain defensive actions (e.g., making head movements more slowly than they otherwise would) that result in a somewhat unnatural behavior. Transfer-of-training studies are therefore in order to determine the extent to which training is impaired by allowing trainees to adapt to simulation environments that tend to produce sickness.

6.1.1.5 Presence

"Presence" is often claimed to be a characteristic of virtual environments. There is no standard definition of presence, but for purposes of this report the term is used to refer to the or trainee's sense of actually being immersed in the situation being simulated. The concept of presence could be a value if it could be shown, as some researchers have proposed, that enhancing presence improves training efficiency and transfer (e.g., Sheridan, 1992).

The recommended research involving presence is composed of three major activities. First and foremost, the concept of presence needs to be operationally defined and a means for measuring it objectively developed. Since presence is a subjective sensation, subjective measures, such as rating scales, are likely to be the primary measurement tool. It may be possible to develop objective measures as well.

Once presence can be measured, the factors that affect it will need to be identified. Situations which have been proposed to enhance presence include: high visual display resolution and large field of view; consistency of information across sensory modalities (including the absence of time delays); absence of artifactual cues; and consistency of actions and consequences between the real and simulated worlds (Held and Durlach, 1992).

The final activity required is to investigate the effect of changes in presence on training effectiveness and transfer. To what extent can increase in presence improve training effectiveness or transfer?

6.1.2 VE Technology

Durlach et al. (1992) suggests a number of research areas concerning the application of virtual environment technology to training. We outline below some areas of study particularly relevant to ICS training of the DI.

6.1.2.1 Visual Systems

No visual cuing device currently matches real-world scene cues with full physical fidelity. All devices will to some extent be limited in terms of resolution, distortion, delay, and noise, and may also exhibit attributes specific to visual display systems such as shimmer. Perhaps the most important topic to be addressed is the degree of visual scene fidelity needed for both training and performance fidelity.

We therefore recommend that a series of experiments be designed to explore fidelity requirements in terms of key performance parameters to determine limiting values (e.g., maximum delay) that allow acceptable task performance and transfer-of-training. Data should be obtained for a variety of tasks, as we expect limiting values to be task dependent. (For example, the visual resolution capability required to identify a large building at 100 meters is likely to be less stringent than the capability needed to replicate the viewing environment when the soldier is hiding in the underbrush. Fidelity requirements may also be object dependent (e.g., greater fidelity for a military target than for a background object such as a tree).

6.1.2.2 Haptic Interface

Two issues need to be addressed with regard to the haptic interface: (1) for which tasks is the haptic interface really needed, and (2) given the need, what are the fidelity requirements?

The answer to the first question depends strongly on the mode of operation. If the ICS allows the soldier relatively free movement, a HMD will be required to provide the visual display, and a programmable general-purpose device will be required to provide a virtual world of controls and other physical objects with which the trainee might interact. But if the trainee's movements are to be relatively constrained, as is most likely to be the case for the near and medium term, the physical interaction with virtual objects becomes optional in many cases.

We suggest transfer-of-training studies to determine the value of simulating physical interaction with objects. For example, we have assumed it would be helpful for CPT and MST to provide the trainee the illusion of actually handling, say, a crew-served weapon, as opposed to manipulating a control switch

or button that causes the simulation of the effects (both in terms of time elapsed and potential damage to the enemy) of using the weapon. Should a transfer-of-training study indicate no loss of training effectiveness with the switch-entry mode, the use of more sophisticated haptic interfaces for this task would be contraindicated. (Recall, the intended purpose of the ICS is to teach primarily cognitive and procedural skills, not basic skills such as weapons handling.)

The field of haptic interfacing is undergoing a period of rapid development that includes introduction of new devices. For interfaces that are considered necessary or desirable for ICS training, a series of studies similar to that suggested above for visual systems is recommended for systems associated with the haptic interface, both in terms of information presented to the soldier as well as position and movement information to be sensed and provided to the simulator. Design requirements with regard to parameters such as resolution, bandwidth, dynamic range, and noise need to be determined for providing acceptable performance and training fidelity. Such studies are especially needed for the more promising new developments.

6.1.2.3 Whole-body Motion Capability

The two issues to be addressed are: is it necessary to provide actual or illusory whole-body motion capability, and, if so, what is the best way to provide it? The need to provide motion capability may be dictated by the outcome of the studies recommended in Section 6.1.2 should it be determined that HMD displays are required or highly desirable, and that whole-body motion capability is required to alleviate simulator sickness by reducing or eliminating visual/vestibular sensory cue conflicts. Barring the need to reduce sickness, motion capability may still be desired to enhance training effectiveness.

There are two aspects to the inclusion of whole-body motion capability within the ICS: the ability for the human to indicate motions to the simulator, and the capability of the simulator to provide acceleration forces or other stimuli to provide the sensation of motion to the human.

Indicating motion to the simulator appears to be the simpler of the two and is likely to be accommodated by devices such as treadmills or large trackballs which allow trainees to move their legs without significantly moving their bodies. Inducing a sensation of motion without actually allowing large body displacements is a more difficult task. Considerable research is needed to develop operational systems to handle whole-body motion in the ICS.

Ideally, research would be first be performed to determine the training benefits of whole-body motion capability, and then

conducted to determine the best way to provide this capability. Research will have to be conducted in the reverse order, however, because the capability must first exist before one can explore its training effectiveness. Should simulator sickness remain an important issue, research will be especially needed to determine what sort of cuing is adequate to overcome the (presumed) sensory conflict. It is not obvious that stimulation short of actual whole-body motion can significantly reduce the sensory conflict arising from a relatively stationary body and visually-perceived self-acceleration.

6.1.2.4 Sensorimotor Inconsistencies

"Sensorimotor inconsistencies" refers to discrepancies between images provided by the visual and haptic systems; i.e., the difference between where the hand appears visually and where it is felt to be. This issue arises when a HMD is used to provide all visual information, including a visual image of the hand. A discrepancy of this type will occur, for example, when the visual system shows an object to be touched or manipulated to be in one location, and an exoskeleton "shows" the object to be in a different location. Research is needed to determine how much of a discrepancy is noticeable by the human and, more to the point, how much of a discrepancy can be tolerated without degrading performance or training fidelity for specific tasks.

6.1.2.5 Computer Generation of Virtual Worlds

The research areas suggested so far have been concerned with the soldier/simulator interface. Research also needs to be conducted into the computational requirements and techniques of providing a virtual world. Areas of interest include:

<u>Acoustic images</u>. Research in the area of acoustic imaging involves both the development of models of the various sound sources (e.g., weapons firing), and algorithms for determining the appropriate signal to be supplied to the various acoustic display devices (e.g. the two earphones or the individual speakers in the speaker array).

<u>Tactual images</u>. Appropriate use of haptic displays requires the development of adequate models of mechanical interactions between objects. Development of such models is highly challenging and as yet in its infancy. Such models may range from simple look-up tables to sophisticated models based on partial differential equations.

<u>Terrain models</u>. Current combat simulations, as implemented in SIMNET, use a pre-computed non-changing terrain database that is replicated in each individual simulator. To provide a more realistic combat environment, the system must be extended to include deformable terrain, i.e., terrain that

is dynamically modified over the course of combat (e.g., cratering, digging a trench). Computational techniques need to be developed to allow the terrain to be modified in a manner that appears realistic to the observer. Algorithms need to be developed for communicating changes in terrain features to the individual simulators connected to the network.

Computational models for human agents. As more degrees of freedom are added to the iconic display of a DI, more sophisticated computational techniques will be needed to direct the movements of these icons so that the icons behave in a manner that is consistent with the terrain database, interact properly with other objects, and otherwise appear to move in a natural manner (e.g., realistic sequence of limb motions when walking or reaching).

See Durlach et al. (1992) for development of these topics.

6.2 Facility Requirements

Independently of the Army's interests in combined arms training, development of VET of potential relevance to ICS training will be done at numerous sites, mostly by private industry. We suggest that a state-of-the-art research facility be developed for the Army - either on-site or at a contractor's location - to provide a centralized location where the effectiveness of VETT for ICS training can be evaluated. This facility will be evolutionary in that capabilities will be built up as the various VE technologies are developed and refined.

We suggest that the proposed research capability be co-located at a facility that is scheduled to obtain an implementation of the CCTT to be developed over the next three years, or, if that is not feasible, that a separate training research facility obtain its own copy of the trainer. This simulator will initially contain "Level 1" technology as defined in chapter 3 of this report and will thus be suitable for use in initial transfer of training studies comparing Level 1 ICS training to current methods.

An advanced HMD is being developed for the Army by CAE Electronics. This device should be available in the same time frame as the CCTT, which would allow the proposed facility to have some Level 2 or 3 capabilities at the outset. Haptic and motion-cuing devices, along with advanced acoustic display systems, should be added to the facility as they become available.

We suggest that the training research facility maintain at least one "high-fidelity" ICS port that includes the most sophisticated VE technology available to the research center.

This port would be the primary DI test bed. Two or three additional "low-fidelity" ports, perhaps containing only Level 1 or 3 capabilities, are suggested to provide DI units that can interact with the primary high-fidelity port for maneuvers that require coordination across multiple DI units.

7. Summary

The applicability of VET to Army training needs was explored. The major goals of this study were to (1) determine projected trends in capabilities and uses of individual combat simulations (ICS), with emphasis on training dismounted infantry; (2) forecast the opportunities and problems associated with using VET for this type of training; (3) identify major hardware and software requirements to allow effective utilization of this technology; and (4) specify research tasks and facility requirements necessary to support ICS research.

7.1 Projected Trends and Capabilities

The Army has issued a procurement for a battlefield simulator known as the Close Combat Tactical Trainer (CCTT). This device will constitute an element of the Combined Arms Tactical Trainer that will integrate infantry, tank crews, and air crews in a combat training environment. A component of the CCTT is the Individual Combat Simulator (ICS) that will allow the individual dismounted infantryman (DI) to interface with the battlefield simulation.

The initial implementation of the CCTT - the first delivery of which is anticipated in 1995 - will accommodate three categories of dismounted infantry: platoon leader, forward observer, and squad leader. The squad leader will control two computer-generated models (icons), each of which represents one of the two fire teams constituting the squad. Comparable personnel will be accommodated when the CCTT operates in the dismounted scout mode.

Preplanned program improvements for the CCTT call for the eventual portrayal of six icons per squad - one for each squad member. Most likely, this will be accomplished by having one trainee playing the role of the squad leader and exercising control over their own icon, with the computer generating control over the remaining five icons. A possible further development might extend the simulator capabilities to allow (live) trainees playing the roles of all squad members. Various DI specialties are expected to be accommodated by the CCTT DI module. It is anticipated that only tank and infantry units will train initially in the combined arms context, with air support and air defense training capabilities to be added later.

The battlefield simulation is expected to accommodate three broad training purposes: (1) combat proficiency training (CPT) to train combat units in the execution of various tactical missions, (2) mission planning and rehearsal (MPR) to develop and refine battle plans and to provide initial training in execution of these plans, and (3) mission-specific training (MST) to provide practice in executing specific missions.

7.2 Opportunities for Virtual Environment Training Technology

The electronic battlefield simulation will incorporate what is broadly known as virtual environment technology, with increasingly complex and sophisticated applications as the technology matures. A three-level progression of technological development is defined. For the most part, Level 1 VET consists of what is required for the CCTT and is largely available now, Level 2 is anticipated to be available in the 3-5 year time frame at reasonable cost, and Level 3 reflects more sophisticated (or problematical) technologies that will likely take longer to become feasible for ICS application.

Level 1 technology includes:

Multi-screen visual displays

Battlefield sounds provided by speakers

Joysticks and other standard control devices to serve as the haptic interface to the ICS

DI icons that portray orientation and stance (standing, kneeling, prone) but do not show articulation of head and limbs

Indication of the general status of the DI (killed, wounded, or active)

Excluded from this level of technology are (1) automated speech recognition, (2) tactile and other haptic displays, (3) sensing of body position, (4) accommodation of whole-body movement, and (5) consideration of the physical condition of the DI other than killed, wounded or active.

Level 2 technology replaces or augments Level 1 technology with:

Visual information provided by low-resolution head-mounted displays (HMD)

Electro-magnetic, electro-optical, or mechanical sensing of limb and body position

Limited speech recognition

Programmable specialized control devices (e.g., joysticks with programmable force/displacement characteristics) to serve as haptic interfaces

Simulation of large-volume movement through movement in place

Articulation of the head and limbs of DI icons

Reflection of the (computed) physical degradation of the DI through limits on the allowable movement speed

Finally, Level 3 technology improves upon (or augments) Level 2 through:

High-resolution HMD

Measurement of eye position if required by the high-resolution HMD

Advanced speech recognition technology

Programmable general-purpose haptic interfaces (e.g., exoskeletons)

Sensory stimulation of whole-body motion involving no actual movement

Fully animated DI models

Modification of the appearance of the DI icon to reflect various states of physical degradation

Application of artificial stress to the trainee to indicate the trainee's (computed) physical state

Level 1 technology is considered adequate for most CPT and MPR training requirements. More sophisticated visual display technology will be required to train in situations that provide rapidly-unfolding activities in close proximity to the soldier (operations in urban and other close-in environments), where the resulting large and rapid head movements place a higher demand on the VE technology needed to supply the visual information adequate for training. An advanced haptic display interface is required for tasks that rely on the sense of touch (such as assessing the surface condition for supporting heavy equipment), or for situations where it is deemed important to have realistic tactile feedback from control and manipulation of weapons, equipment, and other physical objects.

In addition to the requirements stated above, it is recommended that advanced auditory display technology, speech recognition, whole-body movement, and more sophisticated indications of the physical state of the DI be provided for MST, where it is more important to provide a realistic simulation of the battle environment.

7.3 Hardware and Software Requirements

Rapid advances are being made in the areas of visual and auditory displays and in speech recognition technology. Level 2 speech recognition technology is available at present, and Level 2 visual and auditory display technologies are expected to be available at reasonable cost within 3-5 years. Programmable control devices are being refined and are also expected to be feasible for ICS application in the near term, as are tactile displays with limited resolution.

High-resolution tactile devices, general-purpose exoskeletons, and whole-body motion devices are more problematical. Because they pose significant technological challenges, it is not meaningful to forecast a time when such devices will become available for ICS application.

There are a number of modeling issues associated with VETT, some of which require the development of conceptual models (e.g., algorithms for articulated DI representations), and others that primarily require more computational capabilities. Developments along these lines are expected to allow modeling requirements to be met in the 3-5 year time frame.

Physically stressing the trainee to reflect stress that would be incurred in the actual battle environment appears to be an untested concept with respect to combat training in a simulator; there is therefore no basis for predicting the difficulty or rate of development of this capability.

7.4 Recommended Research Tasks and Facilities

Four areas of study are proposed relating to training methodology: (1) methods for measuring transfer of training, (2) training effectiveness of Level 1 technology, (3) enhanced or artificial cuing, and (4) simulator sickness. The latter is a particularly important topic with regard to the feasibility of applying virtual environment technology, because the tendency toward simulator sickness is likely to increase as the visual display is made more compelling.

Six areas of research are also suggested for specific VE technologies and applications:

Explore fidelity requirements for visual displays as a function of the type of object displayed and the training task.

Determine tasks for which haptic interfaces are required, and determine fidelity requirements.

Determine the training benefits to be expected from whole-body motion cuing, and the best way to provide this type of cuing.

Determine allowable discrepancies between visual and haptic display of physical objects.

Explore computational requirements associated with providing a virtual world. Specific areas include acoustic imaging, tactual images, terrain modeling, and computational models for human agents.

Develop a means of measuring presence and identify the factors that affect it.

An implementation of the CCTT is recommended as the core of a research facility to be used for ICS research. This facility should contain at least one high-fidelity ICS port to allow testing of the latest virtual environment technology. This port will serve as the primary DI testbed. From two to four additional low-fidelity ICS ports are suggested to allow the primary DI station to interact with other DI units when appropriate.

8. References

Department of the Army (1988). Mission Training Plan for the Infantry Rifle Platoon and Squad (ARTEP 7-8-MTP). Washington, D.C.

Department of the Navy (1992). Close Combat Tactical Trainer (Solicitation No. N61339-92-R-001). Orlando, FL: Naval Training Systems Center.

Durlach, N. L., Pew, R. W., Aviles, W. A., DiZio, P. A., & Zeltzer, D. L. (1992). Virtual Environment Technology for Training (VETT) (BBN Report No. 7661). Cambridge, MA: Bolt Beranek and Newman Inc.

Held, R. M., & Durlach, N. L. (1992). Telepresence. PRESENCE: Teleoperators and Virtual Environments, 1, 109-112.

Sheridan, T. B. (1992). Musings on telepresence and virtual presence. PRESENCE: Teleoperators and Virtual Environments, 1, 120-125.

9. List of Acronyms

ARI	Army Research Institute
ARTEP	Army Training and Evaluation Program
CCTT	Close Combat Tactical Trainer
CIG	Computer Image Generator
CPT	Combat Proficiency Training
CRT	Cathode Ray Tube
DI	Dismounted Infantry, Dismounted Infantryman
DIS	Distributed Interactive Simulation
FOV	Field of View
HRTF	Head Related Transfer Function
HMD	Helmet-Mounted Display
ICS	Individual Combat Simulation
MOU	Memorandum of Understanding
MPR	Mission Planning and Rehearsal
MST	Mission-Specific Training
NBC	Nuclear, Biological, and Chemical
NTSC	Naval Training Systems Center
PM TRADE	Project Manager, Training Devices
POW	Prisoner of War
SIMNET	Simulation Networking
SOF	Special Operations Forces
STRICOM	Simulation, Training, and Instrumentation Command
VE	Virtual Environment
VET	Virtual Environment Technology
VETT	Virtual Environment Training Technology